# 1989 Annual Report

## Division of Computer Research and Technology

U.S. Department of Health and Human Services

Public Health Service
National Institutes of Health
Bethesda, Maryland 20892

# 1989 Annual Report



## Division of Computer Research and Technology

*U.S. Department of Health and Human Services*

*Public Health Service*
*National Institutes of Health*
*Bethesda, Maryland 20892*

*The Division of Computer Research and Technology (DCRT) has primary responsibility for incorporating the power of modern computers into the biomedical programs and administrative procedures of NIH. DCRT serves as a scientific and technological resource for other parts of the Public Health Service, and for other Federal Organizations with biomedical and statistical computing needs.*

*DCRT programs focus on three primary activities: conducting research, developing computer systems, and providing computer facilities.*

*The* ***Office of the Director*** *provides overall program direction for DCRT, and serves as a central NIH focus for automated data processing policy. In addition, the Office sponsors research and development work in molecular graphics and simulation.*

*The* ***Computer Center Branch*** *designs, implements, and operates the NIH Computer Center and provides assistance, training, and technical communications to the nearly 17,000 users.*

*The* ***Personal Computing Branch*** *provides guidance and support to scientists and administrators throughout NIH in the effective use of personal workstations, local area networks and associated automation technology. PCB works closely with other DCRT labs and branches to monitor these rapidly changing technologies with a view toward the technical requirements of the user community so that these needs are accommodated in the development and support provided by DCRT.*

*The* ***Data Management Branch*** *serves as central systems analysis, design and programming resource for data processing projects relating to scientific, technical, management, and administrative data.*

*The* ***Laboratory of Applied Studies*** *relates mathematics, statistics, and computer sciences to such biomedical problems as ECG analysis, evaluation of physiological systems in health and disease, and estimation problems in laboratory medicine.*

*The* ***Laboratory of Statistical and Mathematical Methodology*** *provides statistical and mathematical help in the computer analysis of biomedical data and offers statistical and mathematical packages for users.*

*The* ***Physical Sciences Laboratory*** *conducts research in mathematical theory and practical instrumentation to explain biological phenomena in terms of chemistry and physics at the subcellular molecular levels.*

*The* ***Computer Systems Laboratory*** *provides consultation and collaboration in the design and implementation of specialized computer systems for laboratory and clinical applications.*

# From the Director

*The year 1989 marked the 25-year anniversary of DCRT, and saw major advances in the computing facilities available to the NIH staff. These advances include increasing and updating mainframes, providing a minisupercomputer for scientific computing, developing a users cadre of highly advanced graphics workstations for NIH molecular biologists, and facilitating scientific machine-based communication both on campus and globally.*

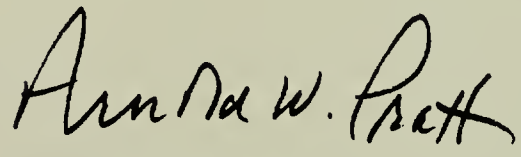
*A major 10-year contract was awarded to IBM that ensures that NIH will have access to state-of-the-art mainframe computing technology through the 1990's. Another contract, to the Convex Computer Corporation, will provide a minisupercomputing capability to replace DCRT's long-reliable but obsolete DECSystem-10. The Convex will enhance scientific computing at NIH by offering applications that include programming for analysis of protein and nucleic acid structure, communications, and integrated graphics.*

*The division also led a project jointly funded by DCRT and the Institutes to define and purchase 40 graphics workstations and make them available to NIH scientists campus wide. These machines will provide chemists, molecular biologists, and geneticists with powerful research tools for computational science.*

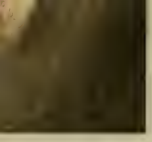
*A Network Task Group was established to provide the technical, functional, and support services for a campus-wide communications structure. This new organization has three main functions: to establish a campus area network; to provide support and guidance to local area networks presently operating on campus; and to provide support for the DCRT network. Additionally, DCRT helped to establish an NIH Network Policy Board, consisting of several senior NIH managers, who will provide policy and institutional guidance for networks throughout NIH.*

*A major DCRT commitment to the "Administrative Data Base" continues with award of a contract for support and maintenance of this huge software package that facilitates more than 75,000 transactions a day, spanning procurement to travel. DCRT staff will eventually be relieved of this commitment and be better able to pursue the re-engineering of the system.*

*These initiatives are part of an ongoing process to meet the ever changing challenge of keeping NIH and the biomedical community at the forefront of computer technology.*



*Arnold W. Pratt, M.D.*
*Director, DCRT*



# Contents

# Office of the Director

*Arnold W. Pratt, M.D., Director*

The Office of the Director provides overall program and management direction for DCRT. The Director, Deputy Director, Assistant Director and Executive Officer work together as the immediate Office of the Director.

Two other offices supplement the work of the DCRT laboratories and branches:

The **Office of Scientific and Technical Communication** (OSTC) serves as central source of information about DCRT activities and about computing and related disciplines. It includes the DCRT Information Office and the DCRT Library.

The **Office of Administrative Management** (OAM) provides administrative and managerial support for the work of DCRT. OAM includes the Administrative, Personnel, Financial Management and Project Control Offices.

Also within the Office of the Director is the **Equal Employment Opportunity Office** (EEO), which manages a full EEO program for the Division. The office serves as the focal point and advisory for all activities relating to the equal employment opportunities of DCRT employees and applicants. The EEO Officer maintains a close working relationship with the NIH Division of Equal Opportunity and other components concerned with minority and women's issues.

The division's role in NIH Information Resources Management (IRM) activities is shared among members of the immediate Office of the Director, but responsibility for NIH ADP planning, budgeting, and security still resides in the ADP Policy section.

The Office of the Director also has a molecular graphics and simulation research component that applies computing technology to biomedical research, to solve problems in macromolecular structure representation and modeling for collaborating NIH scientists and visitors.

## Network Task Group

*David C. Songco, Chief*

A Network Task Group (NTG) was established in the Office of the Director in May 1989 to bring increased emphasis on DCRT networking activities. This group brings together, under one manager, DCRT staff responsible for networking development, support, and guidance at NIH. Staff members were detailed from the Computer Systems Laboratory. NTG consists of seven engineers and one computer specialist.

Primary responsibility for NTG is the design and implementation of a "backbone" network to link locally managed networks on campus. In recognition of the complexities of implementing an NIH-wide facility of this magnitude, DCRT recommended that NIH form a Network Policy Board, with a charter to oversee the engineering, installation, management, administration of, and services provided by the campus backbone network.

The first task of the NTG was to develop a status report on the campus backbone project, including accomplishments, problems, schedules, and budgets. This report, presented to the Network Policy Board on August 17, 1989, described increased estimated costs and substantial schedule delays and uncertainty. NTG described several possible midcourse corrections to offset the projected increases in the cost of network implementation. These alternatives required additional study and will be presented in early FY90.

In addition NTG worked closely with the Personal Computing Branch to provide guidance and support for locally managed networks and developed and supported DCRT inhouse networking activities in collaboration with the Computer Systems Laboratory and the Computer Center Branch.

## Office of Administrative Management

*Marian Dawson, Chief*

The Office of Administrative Management (OAM) serves DCRT by providing administrative support and resources allocation. Program planning and evaluation and policy and legislative analysis functions also reside in OAM. The office is organized functionally into three sections: finance, personnel, and administration.

The Administrative Section provides primary support for all administrative activities within the division. The Personnel Management Section provides a full range of personnel services to the more than 300 DCRT employees. The Financial Management Section provides budget formulation and execution functions as well as management of the project accounting system (PAS), which provides detailed billing information to users of fee-for-service computer and programing services provided by the division.

During the past year OAM staff continued their involvement in several major procurements. Much time was devoted to the procurement of the DECsystem-10 replacement, resulting in a contract award on May 15, 1989. In addition, staff assistance was provided for the NIH-wide procurement of molecular graphics workstations for distribution throughout the NIH Intramural community.

As part of the continuing efforts to improve the procurement process at NIH, DCRT responded to a request to conduct a study within the division. The study resulted in a proposal to

streamline the DELPRO ordering process by consolidating ordering activities carried out by DCRT ordering officials. The new procedure will be implemented in early FY90.

The DCRT Personnel Management Section implemented the new Employee Performance Management System's (EPMS) Awards Program. This program complements the SES and PMRS program to provide annual performance-based awards to virtually all DCRT employees. The EPMS Awards Program was successful due to the communication, training, and coordination efforts of the personnel office staff.

DCRT recruitment activity continues at an active pace aided by a number of special hiring authorities, particularly for entry level professionals. Personnel office staff attended several career fairs and contacted numerous area colleges and technical schools in an attempt to attract new employees. Once again, DCRT was able to augment staff by using the summer program. Approximately 25 university and high school students were hired for technical, professional, and clerical positions throughout the division.

The staff continued to provide a wide variety of assistance and information on employee benefits and services including the new leave-sharing program and the new awards programs through follow-up interviews, memos, status reports, and formal and informal meetings with employees. Staff attended specialized OPM training courses and professional conferences. The office also expanded its computer capabilities by adding a pc-based automated personnel information system and a CD ROM version of the Federal Personnel Manual (FPM).

The administrative office devoted much effort this year to the division's procurement process. All staff participated in the latest DELPRO training class offered by the Division of Procurement, and in turn, provided briefing sessions to all Division personnel involved or interested in the procurement process. The sessions were well attended by staff. The administrative office carried out the lead review responsibility for the DELPRO study. Efforts continued on the refinement of the procurement status report for DCRT.

The administrative staff participated in the pilot study for the online processing of travel through the Administrative Data Base system, with excellent results. The staff also prepared to implement a credit card system to be used in lieu of travel advances for the frequent traveler.

The budget office continued to carry out annual budget functions for DCRT. In addition, a variety of monthly reports were furnished to program managers to inform them of their budgetary status. The office also participated in the annual rate-setting process for Divisional Service and Supply Fund activities, and the ITS Budget Submission.

Staff of the project control office processed requests for new accounts and users of data processing activities, opening 134 new accounts and registering 2,604 new users. The office also completed its annual update of information on over 3,300 accounts and over 19,200 users currently in the project accounting system.

## DCRT Information Office

*Colleen Henrichsen, Chief*

The DCRT Information Office is responsible for a comprehensive information program directed toward a variety of audiences within and outside NIH. The program is addressed to those whose

needs and attitudes have important bearing upon what DCRT strives to accomplish.

The information office continued its involvement in many events and activities affecting the division, the NIH community, and the general public in FY89.

It was a busy year for press involvement with inquiries concerning DCRT programs and activities including computers in biomedical research, molecular modeling, image processing, and advanced laboratory workstations. These requests came from a number of general interest publications including the *Washington Post,* the *Washington Times,* the *New York Times,* the *Baltimore Sun,* and the *Chicago Tribune;* as well science, computer, and professional publications including *Government Computer News, Federal Computer Week, Science News, Omni, ComputerWorld, Datamation Magazine, Personal Computing Magazine, Electronic Engineering Times, Communication Week,* and *MIS Week.*

Information office staff created, developed, and produced a new publication, *DCRT Computer Clips,* a weekly compilation of articles on computer technology advances as reported in the popular press, which is distributed to DCRT staff.

In addition, the information office responded to more than 1,500 telephone inquiries and distributed nearly 10,000 publications on DCRT and NIH programs. As part of its public relations effort, the office placed more than 20 stories in the *NIH Record.*

To provide the NIH community with current information about DCRT programs, the office produced new editions of *Computing Resources,* the *DCRT Annual Report,* and the *Directory of NIH Image Processing Facilities.*

In an effort to attract high-quality, engineering and computer science graduates to DCRT, the information office produced a full-color recruitment brochure, which was distributed to nearly 100 minority colleges and universities. More extensive mailings of the publication are planned for FY90.

NIH publications contributed to by the DCRT information office in FY89 include the *Report of International Activities, Research Advances, NIH Almanac, NIH Information Index, NIH Scientific Directory—Annual Bibliography,* and *NIH Publications List.*

New, sophisticated color presentation graphics were developed and produced for the DCRT Director. These, along with other information office-produced slides were presented to various groups including the NIH Scientific Directors and corporate groups.

The staff also facilitated the development, script writing, and production of a videotape on computer security awareness for the ADP Policy Coordination Office. The videotape will be distributed to all BID's to be shown to NIH employees who use computers as part of an extensive computer security awareness effort.

The office facilitated tours of the DCRT computer facility, hosting various national and foreign scientists as well as professional organizations.

In other information related activities, the staff participated in Career Day, sponsored by the NIH Women's Advisory Council; coordinated the Savings Bond Drive for DCRT; and participated in the Division coordinated Combined Federal Campaign. The office also coordinated the annual DCRT picnic.

Information office staff honed their skills and learned new ones with training in a variety of areas including computer technology, management and supervision, and editing.

## DCRT Library

*Ellen Chu, Chief*

The DCRT Library maintains a special collection in computer science, mathematics, and statistics, along with computer applications in biomedical sciences, engineering, information science and management. Its information resources and services support DCRT as well as other parts of NIH. Information retrieval and referral services include online search services for citations or full-text data, acquisition or borrowing of publications, and referrals to other information sources.

Fifty-five percent of the registered borrowers work in other parts of NIH. This year library staff processed acquisition or interlibrary borrowing requests of 1,300 books, reports, articles, journal issues, and subscriptions. Once again the DCRT Library lent out more books to other libraries than it borrowed, continuing a pattern started in 1978. The staff continued implementation of the network online catalog. It installed and operated its own independent local area network (LAN) providing DCRT staff network access to the online catalog, library information files, and electronic mail to communicate and transmit data and news.

The library installed additional compact disk (CD) information retrieval systems and implemented direct end-user searching of MEDLINE and related files. The number of online searches performed by the library staff decreased from the previous year, reflecting the trend toward end-user searching. (Extensive searches were conducted for DCRT staff working on major procurements.) As users shift to executing their own computerized searches, the staff has begun to provide training and advice on effective searching techniques. This year, a project with Personal Computer Branch staff investigated network access to CD's, culminating in the selection and installation of a system for multi-user DCRT LAN access. Copyright and licensing arrangements will determine if CD's will be installed on the network or at a single library workstation. The library project in network information delivery systems will continue with investigation of Macintosh access.

The year 1989 saw a major rearrangement of the library. An open house featured a demonstration of new information systems in reconfigured staff and user spaces to accommodate LAN hardware and personal workstations. Library staff demonstrated the network online catalog; CD editions of *Books in Print, Oxford English Dictionary, Science Citation Index,* and *Computer Library* (a full text database); and remote access to search the NIH Library catalog. Demonstrations of the JURIS system were provided to administrative staff. Librarians from universities, associations, the federal government, and Germany visited the Library to see various computerized systems. The library cosponsored the organization of a Washington area users group for one library application system and hosted the first meeting at DCRT in October.

The library chief completed the first year of a three year term on the Special Libraries Association Scholarship Committee. She completed her one year term chairing the Finance Committee of the D.C. Chapter.

## Research Projects

### Information Retrieval Systems on Local Area Networks

*Ellen Moy Chu*
*with Clark C. Collins (DCRT/PCB)*

The project objective is to develop local area network (LAN) access to DCRT Library information systems. Previous work included installation of a LAN server and development of LAN access to the library online catalog. This year, an independent LAN was installed, with continued LAN access to library online catalog and information files. Demonstrations were provided to staff of the NIH Library and the NCI Documentation Center. An extensive feasibility survey of network compact disk (CD) systems was conducted. With Clark Collins of the Personal Computer Branch, software was installed for multiuser network access to CD information systems. CD's provide user access to very large text and full-text databases for information retrieval. Future plans include testing CD systems for Macintosh workstations. Multimedia CD's with audio and video capabilities will be investigated as they become available. Evaluations of specific CD products will continue as each producer provides its own search engines and system designs. Progress to date has enabled the DCRT Library to provide electronic information delivery and extended Library access to remote workstations on the DCRT LAN. The knowledge and experience will provide the basis of advice given to other NIH organizations.

## Policy Coordination Section

*John M. Campbell, Chief*

The Policy Coordination Section has responsibility for ADP policy coordination and reporting functions for NIH. These include ADP procurement reviews, the information technology systems budget, ADP annual planning, administrative systems, inventories and hardware, Federal Information Processing standards (FIPS), and ADP systems security. The section provides staff support to the Director, DCRT, and the Associate Director for Administration, NIH, for Information Research Management (IRM) activities.

A major emphasis during FY89 was refinement and dissemination of the automated Standard Risk Protocol for completing risk assessments and risk analysis covered by the Computer Security Act of 1987. This personal computer application consolidates the requirements of OMB Circulars A123, A127, and A130, and the requirements imposed by chapter 2 of the DHHS IRM Manual. This unique automated process responds to the specific criteria covered by the review processes and is designed to measure the level of security in an application system/facility, and to provide management with guidelines for implementing cost effective safeguards to improve the security risk level.

A system review module is designed to determine the effectiveness of a system's development and operation, and to determine the effectiveness of the output reports generated by the system. The module examines the level of compliance with GAO principles and standards on the effectiveness, reliability, and integrity of financial applications.

A summary risk assessment program collects and stores data relative to risk analyses, contingency plans, reviews, and certifications. Cross-summarization of the data relative to systems, facilities, and equipment, provides management with the capability of correcting vulnerabilities through summary analyses.

A contingency planning module will be developed to characterize the wide and diversified range of requirements involved with NIH ADP activity. The module will provide the capability of interaction among facilities or other sources to recover from adverse events that prevent the facilities from providing day-to-day ADP services to the users. A contingency statistical module will calculate within different scenarios the chances of incurring adversities and relate these possibilities to the dollars required to recover from different types of disasters. The statistical base relates the most cost-effective solutions to different disasters.

A virus module will look to the problems that facilities have faced with illegal attempts to penetrate operating computer systems and will provide a tool to measure the vulnerabilities of operating systems.

The Policy Coordination Section has published a user's manual that contains three diskettes and is currently in use at NIH and other PHS Agencies. The Protocol has been accepted as a cost-effective approach to managing ADP security.

## Molecular Graphics & Simulation

In FY89 the molecular graphics laboratory, led by Richard J. Feldmann, participated in the acquisition of more than 40 high- and low-level Silicon Graphics workstations equipped with molecular modeling software for scientists located NIH wide. The workstations are equipped to run four packages: two computer-aided molecular design (CAMD) packages from Polygen—Quantum and Charm; a molecular display program called GEMM developed by B. K. Lee of DCRT; and Sequence Editor Aligner (SEA), developed by Ann Barber at the Frederick Cancer Research Center.

This purchase comes as a result of more than 15 years of molecular graphics system development in DCRT and usage by NIH scientists. During these years molecular modeling by computer has become of increasing interest and use to NIH scientists. However, because the large and expensive equipment needed for this function was located in the molecular graphics laboratory, scientists were required to bring their work to the equipment, located in Building 12A, to investigate molecular structure problems and receive collaborative support from the staff. The scientists then had to return to their laboratories to examine their data and prepare for the next session.

The demand for molecular structure modeling grew even more as a variety of proteins, protein fragments, and nucleic acids of the viruses associated with AIDS began to be modeled. As a result, increasing numbers of scientists relied on the molecular graphics laboratory for support.

While demand for sophisticated graphics modeling grew the price of graphic workstations was coming down, making it feasible to install them in individual laboratories, placing critical tools directly in the hands of a community of leading NIH scientists. This newly available alternative is likely to have a significant impact on many research efforts at NIH, including the

advancement of research efforts involving AIDS related viruses.

**Advances in Molecular Simulation**

The fundamental tools to study molecules of biological interest theoretically are molecular dynamics, molecular mechanics, modeling, and ab initio analysis of small molecule structure. Work is proceeding on three fronts: (1) method development, (2) basic research problems, and (3) applied projects. Method development is comprised of two areas of interest. They are the characterization of new theoretical techniques, and the redesign of existing methods to fully use special laboratory machines, or for machines being designed by the lab with the hope that they achieve the performance of 20–30 Cray X/MP's. Basic research projects are designed to provide a better understanding of biochemical systems and to find better ways to use existing methods. The applied projects of the molecular graphics laboratory are usually carried out in close collaboration with experimenters from NIH and elsewhere, and often involves the three-dimensional structure determination of macromolecular systems. The laboratory has recently completed three major projects and several others are in progress that use molecular dynamics simulations to predict unknown protein and peptide structures. These projects are applied with specific biomedical goals in mind, such as vaccine development.

## Research Projects

**Theoretical Study of beta-Lactam Antibiotics**

*Richard J. Loncharich, Ph.D., and Bernard R. Brooks, Ph.D.*

The GAUSSIAN 86 series of programs has been used to carry out ab initio quantum mechanical calculations to model nucleophilic additions to beta-lactam antibiotics. So far the staff has studied the attack of the beta-lactam by water. There are two transitions states to the formation of product. In the first water attacks the carbonyl with transfer of one hydrogen to the carbonyl oxygen. The second occurs after ring opening and is a direct transfer of a hydrogen atom to the nitrogen. At the moment, the mechanical results appear to be basis set dependent, however, this project is still in progress and the prediction of a mechanistic pathway remains unanswered. Further aspects of the electronic nature of nucleophilic addition to the beta-lactam will be studied by substituting an amine (donating) group or a formyl (withdrawing) group at the carbon atom adjacent to the carbonyl. Placing substitutents on the lactam ring will give an indication of the polarization changes that occur as nucleophilic addition proceeds. It will also be of interest to study the binding of beta-lactams using molecular dynamics approach for simulating chemical reactions.

**The Effects of Truncating Long-Range Forces on Protein Dynamics**

*Richard J. Loncharich, Ph.D., and Bernard R. Brooks, Ph.D.*

Investigators have carried out a basic research project that considers the effects of truncating long range forces on protein dynamics. Six methods of truncation investigated as a function of cutoff criterion of the long range potentials were: (1) a shifted potential; (2) a switching function; (3) simple atom-atom truncation based on distance; (4) simple atom-atom truncation based on a list that is updated periodically (every 25 steps); (5) simple group-group truncation based on distance; and (6) simple group-

group truncation based on a list that is updated periodically (every 25 steps). Based on seventy calculations of carboxy-myoglobin researchers have shown that the method and distance of long range cutoff have a dramatic effect on overall protein behavior. Evaluation of the different methods is based on comparison of a simulation's rms fluctuation about the average coordinates, the rms deviation from the average coordinates of a no cutoff simulation and from the x-ray structure of the protein. The simulations in which long range forces are truncated by a shifted potential shows large rms deviations for cutoff criteria less than 14 Angstroms, and reasonable deviations and fluctuations at this cutoff distance or larger. Simulations using a switching function were investigated by varying the range over which electrostatic interactions are switched off. Results using a short switching function that switches off the potential over a short range of distances are poor for all cutoff distances. A switching function over a 5–9 Angstrom range gives reasonable results for a distance dependent dielectric, but not using a constant dielectric. Both the atom-atom and group-group truncation methods based on distance shows large rms deviation and fluctuation for short cutoff distances, while for cutoff distances of 11 Angstroms or greater, reasonable results are achieved. Although comparison of these to distance based truncation methods show surprisingly larger rms deviations for the group-group truncation, contrary to simulation studies of aqueous ionic solutions. The results of atom-atom or group-group list based simulations generally appear to be less stable than the distance based simulations, and require more frequent velocity scaling or stronger coupling to a heatbath. Researchers hope that the results of these simulations provide assistance in choosing a method of truncating long range potentials in protein dynamics.

**The Temperature Dependence of Dynamics of Hydrated Myoglobin: Comparisons of Force-Field Calculations with Neutron Scattering and X-Ray Data**

*Richard J. Loncharich, Ph.D., and Bernard R. Brooks, Ph.D.*

Eleven molecular dynamics simulations of 150 psec. in length are carried out on the carboxy-myoglobin protein at 20, 60, 100, 180, 220, 240, 260, 280, 300, 320, and 340 K. The simulations attempt to mimic neutron scattering experiments very closely by including 349 waters around the protein. Theoretically determined elastic, quasielastic, and inelastic neutron scattering data are directly compared with experiment. The elastic scattering is decreased compared to experiment. Inelastic and quasielastic spectra show that the inelastic peak is shifted to lower frequency than the experimental value, while quasielastic behavior is in good agreement with experiment. In addition, researchers analyze simulation trajectories by comparing rms deviations from the starting structure, average atomic fluctuations, and a helical time series. The anisotropy and anharmonicity of the atomic position distributions are also determined. The anisotropy is small below 180 K and dramatically increases above this temperature. The anharmonicity is large at all temperatures. Overall simulation results indicate that the protein behaves with near harmonic vibration in well defined substates at temperatures below 180 K. Protein behavior becomes much more anharmonic with transitions between many conformational substates at higher temperatures.

**Modeling and Simulation of Lipid Crystals and Lipid Bilayers**

*Bernard R. Brooks, Ph.D.; Chris J. Kuhn; Richard J. Loncharich, Ph.D.*
*with Richard W. Pastor (FDA/CBER)*

The long range goal of this project is to be able to accurately simulate membrane bound peptides and proteins. Current work involves the simulation of a small lipid crystal (DPPE) using many different energy parameter sets and comparing the results. The analysis of the theoretical crystal involves calculation of the pressure, energy, atomic fluctuation, and atomic deviation from crystal observables. The results are used to generate better energy parameters. Future work after optimizing the energy parameters will be to model and simulate a lipid bilayer and then model other macromolecules in the membrane environment.

**Conversion of Physical Models into Three-dimensional Coordinates for Computer Analysis and Simulation**

*Stephanie Burton and Bernard R. Brooks, Ph.D.*

Although the staff has been carrying out detailed molecular simulation for several years, and has used plastic molecular models to test and explore ideas, it has never had any convenient way to convert between plastic data and computer data. Investigators are nearing the completion of a small project that allows scientists to present ideas in a plastic model and directly scan such models into a computer using several photographs. The Apple Macintosh II computer is a base for this project due to the availability and ease of connection of peripherals, such as high quality scanners and high resolution frame grabbers (direct scanning camera), to this machine. This software tool should allow researchers to obtain three-dimensional coordinates from stereo pictures in published journal articles or other sources. In the future, this tool will be augmented to allow the use of a color frame grabber and a zoom feature will be added so that large complex structures can be processed with precision.

**The Structure Determination of AIDS Glycoprotein Fragments**

*Bernard R. Brooks, Ph.D.*
*with F.W. Carson (American University); R.M. Venable (FDA/CBER); R.W. Pastor (FDA/CBER)*

Investigators are now actively involved with the intramural effort to develop an antiviral agent for AIDS. Initial published work has focused on the three-dimensional structure determination of selected portions of the surface glycoproteins gp120 and gp41. The work performed to date proposes possible structures for the amphipathic segments of the HIV-1 gp41 envelope protein, and may have implications for current vaccine development efforts. Results suggest a possible mechanism for the cytopathic effects of the gp120-gp41 complex. Researchers await further experimental data from NMR or from x-ray crystallography to support further efforts in designing an antiviral agent for this potential target region.

**Theoretical Analysis of Peptide Binding to the Active Site of HIV-1 Protease Using Molecular Dynamics**

*Bernard R. Brooks, Ph.D.*
*with F.W. Carson (American University); R.M. Venable (FDA/CBER).*

Investigators are currently carrying out a series of molecular dynamics simulation on the HIV-1 protease. In these preliminary studies, researchers are examining the interactions between the substrate and the protein in a solvated system.

In future simulations, investigators will examine in detail, the behavior of the flap regions to try to identify their role in specificity. By varying the substrate, scientists hope to be able to identify the key interactions that account for the specificity of this protease. Researchers anticipate collaborating with a pharmaceutical company so that ideas that derive from the analysis of the simulations can be tested experimentally.

**Harmonic and Langevin Analysis of Large Systems**

*Dusanka Janezic, Ph.D.; Bernard R. Brooks, Ph.D.*
*with R.M. Venable (FDA/CBER)*

In this ongoing project, the use of harmonic analysis of large systems is used in a variety of ways to provide a better understanding of a protein and RNA systems. This work has involved performing many different types of harmonic analysis methods to the bovine pancreatic trypsin inhibitor as a trial system. Both harmonic and quasiharmonic analysis calculations have been carried out in several different basis. These results are currently being analyzed and written up. There has also been extensive analysis of the time behavior of the results from harmonic analysis and a direct comparison with simulation results. Further work involving the Langevin mode analysis of this trial system and other systems is planned. From this work investigators hope to: (1) better explain neutron spectral (time of flight) data; (2) better characterize friction and solvent damping; (3) find stress points in macromolecules; and (4) explain (or predict) some site specific mutation data.

**Development of New Theoretical Methods for Studying Macromolecules**

*Bernard R. Brooks, Ph.D.; and Richard J. Loncharich, Ph.D.*
*with B.K. Lee (DCRT/PSL); R.M. Venable (FDA/CBER)*

One of the most interesting and important areas of ongoing research is that of the development of new methods. This ongoing work involves the development and characterization of new techniques that have not been tried before. Examples of this are intelligent methods for processing conformational searching data, algorithms for performing Langevin normal mode analysis on macromolecules, new methods for performing free energy perturbation calculation, methods for treating solvent implicitly to provide for hydrophobic affects without the explicit inclusion of many water molecules, methods to allow a better use of NMR data in determining solution structures of macromolecules, and methods to rapidly compute the accessible volume and surface areas of a macromolecule. Most of these methods are being developed to provide tools for current needs as well as tools for applied projects in the coming years.

**Theoretical Study of Polycyclic Aromatic Hydrocarbon-diol-epoxide-DNA Adducts**

*Bernard R. Brooks, Ph.D.*
*with A. Weston (NCI)*

In this ongoing study, the binding of diol-epoxides to DNA is studied. Investigators have approached this project by the modeling and simulation of carcinogenic polycyclic aromatic hydrocarbon-diol-epoxide binding to DNA. The two main questions that this work attempts to answer are: (1) Can the wide difference in carcinogenic activity of similar adducts be

explained from a theoretical study? and (2) Can the cross-reactivity data between adducts and antibodies grown against different adducts be explained? The results of this work indicate that there is a strong correlation between carcinogenic activity and orientation of the hydrocarbon in the minor groove of DNA. This work also provides a possible explanation to observed cross-reactivity data of antibodies grown against one adduct and DNA bound to different adducts, which is that the antibodies recognize the distortion of the DNA backbone rather than the adduct itself and that structures that cause similar phosphate group movements have high cross reactivities. This study also suggests that there is a strong DNA sequence dependence for carcinogenic activity of polycyclic aromatic hydrocarbon-diol-epoxides.

## Publications and Presentations


Pastor, R.W.; Brooks B.R.; and Szabo, A.: Algorithms and Error Analysis for Langevin Dynamic Simulations. *Mol. Phys.* 65, 1409–19 1988.

Venable, R.M.; Pastor, R.W.; Brooks, B.R.; Carson, F.W.: Theoretically Determined Three-Dimensional Structures for Amphipathic Segments of the HIV-1 gp41 Envelope Protein. *AIDS Res. and Hum. Ret.* 5–7 1989.

Brooks, B.R.: Molecular Dynamics for Problems in Structural Biology. *Nobel Symposium Proceedings, Springer Lecture Notes,* 1989 (in press).

Loncharich, R.J.; Brooks, B.R.: Long Range Truncation Effects in Protein Molecular Dynamics Simulations. *Proteins* 1989 (in press).

Han, K.H.; Syi, J.L.; Brooks, B.R.; Ferretti, J.A.: Solution Conformations of the B loop Fragments of the Human Transforming Growth Factor-@a and Epidermal Growth Factor by Nuclear Magnetic Resonance and Restrained Molecular Dynamics. *Proc. Natl. Acad. Sci.* 1989 (in press).

Hooper, C.A.; Reid, R.H.; Brooks, B.R.: Computer Simulations of a Tumor Surface Octapeptide Epitope. *Biopolymers* 28, 525–539 1989.

Kay, L.E.; Brooks, B.R.; Sparks, S.W.; Torchia, D.A.; Bax, A.: Measurement of NH-Ca^H Coupling Constants in Staphylococcal Nuclease by Two-Dimensional NMR and Comparison with X-ray Crystallographic Results: *J. Am. Chem. Soc.* 111, 5489 (1989).

# Computer Center Branch

*Joseph D. Naughton, Chief*

The Computer Center Branch designs, implements, operates, and maintains the central NIH Computer Utility and its associated telecommunications facilities. Two interconnected, integrated multicomputer facilities, designed around large-scale IBM and DEC mainframe computers, make up the NIH Computer Utility. The utility has evolved over two decades to provide computational services to meet the dynamic and diverse requirements of NIH research investigators and administrators in the conduct and management of modern biomedical research. The utility provides interactive timesharing, database management, graphics, batch, high performance scientific computation (e.g., vector), and word processing services to approximately 18,000 authorized users at NIH and in 26 agencies throughout the federal government as a Federal Data Processing Center (FDPC). The utility has been augmented to include a Convex minisupercomputer, which will, in time, replace the DEC mainframe system.

The IBM facility, valued at approximately $100 million, is made up of six IBM 3090 processors, each with one or more vector facilities, and an array of supporting peripheral equipment such as disks, printers, magnetic tapes, communications controllers. The 6 processors contain a total of over 768 million bytes of directly addressable memory. The peripheral complex includes 728 online disk drives with a total data storage capacity of over 1.2 trillion bytes; 108 cartridge tape drives with a transfer rate of 3 MB per second; 20 reel type tape drives with a transfer rate of 780 KB per second, and 4 drives with a transfer rate of 200 KB per second; 7 18,000 lines-per-minute page printers, 3 45 page-per-minute duplex cut sheet printers, 2 card reader/punches; and 13 telecommunications controllers, supporting a teleprocessing network of over 1,500 lines of various bandwidths connected to 330 remote job entry (RJE) computers and an estimated 10,000 interactive CRT and/or hardcopy terminals located throughout the world.

The DECsystem-10 facility is designed around three KL10 processors with 2 1/2 million words of main memory; 12 554 megabyte disk drives and 19 197 megabyte disk drives; 7 780 KB/second and 3 200 KB/second tape drives; and 11 communications processors capable of supporting over 75 simultaneous timesharing users. The Convex system consists of a dual processor Convex C220 minisupercomputer with 128 megabytes of memory, 8 gigabytes of disk storage, 4 magnetic tape drives, and 3 I/O processors dedicated to handling input and output from disk drives, tape drives, Ethernet connections, and terminals.

The IBM, Convex, and DEC facilities are linked together by high-speed telecommunications lines to facilitate the exchange of data, electronic mail, and programs, and for printing services. Ancillary equipment includes two computer output microfiche units and film processors, two four-color high resolution plotters, and other miscellaneous devices.

Services of the utility are available worldwide through the TYMNET, BITNET, and Internet international data communications networks, as well as through the Federal Telecommunications Service (FTS), and commercial switched telephone services. Some users and local area networks (LAN) are connected to the utility with dedicated (hardwired) communications lines.

The NIH Computer Utility operates 24 hours a day, 7 days a week, processing over 11,000 interactive sessions, 75,000 database transac-

tions, and 18,000 batch jobs daily. Over 90 percent of all interactive commands are executed with subsecond response time, and 6 service classes for batch jobs guarantee maximum turnaround times of less than 30 minutes, 1 hour, 2 hours, and overnight. All work is processed on a fee-for-service, full cost-recovery basis offering a 60 percent discount for work processed after 5 p.m.

Applications are programmed in FORTRAN, COBOL, PL/1, WYLBUR Command Procedures Language, BASIC, Pascal, C, and Assembly Language, using several interactive programming systems (TSO, WYLBUR, Convex and DECsystem-10 Timesharing). Database management services are provided through DB2, a general purpose, open-shop, relational database system, and IMS, a hierarchial, interactive transaction processing system for administrative applications. A variety of statistical analysis, modeling, data management, and utility programs are available, as well as the TELL-A-GRAF, POSTER, and OMNIGRAPH interactive graphics packages.

An in-house training program offers 39 specialized classroom lecture courses in 2 semesters and a summer term. An online interactive Assisted-By-Computer (ABC) training system permits users to develop hands-on expertise in the use of the utility in their own work environment. A variety of 12 different textbook courses are lent to students for topics ranging from "Introduction to Computer Programming" to "ISPF Dialog Management." Technical documentation provided ranges from a comprehensive Computer Center Users Guide, which defines all available services, standards, and procedures, to a technical newsletter, INTERFACE, which keeps users up to date on the current status of the utility and announces changes and upgrades as they occur.

A highly trained staff of professional, technical, and management personnel, provides assistance to users and ensures smooth functioning of the NIH Computer Utility. Computer specialists, systems programmers, and analysts develop and maintain operating systems software, provide technical consultation on program design and problem resolution, teach training courses, and write technical documentation. Computer system technicians and operations personnel operate and maintain the computer utility hardware and telecommunications networks, and provide user support services. The Computer Center Branch management team develops long-term program goals, conducts acquisitions, assists users with project planning, maintains the design integrity of the utility, and responds to changing user demands by designing and implementing new features and services.

The effectiveness of computers in the support of modern biomedical research is furthered by research and development projects conducted by the Computer Center staff.

## Communications and Connectivity Enhanced

In order to create more powerful tools for biomedical research, the NIH Computer Center has developed a plan for a comprehensive communications system that allows users with a personal computer or laboratory workstation to access the services of the NIH Computer Utility using either a local area network (LAN) or dial-up connections.

The Computer Center began in the late 1960's with a primary orientation toward access to all

its services by all its users at their work sites over conventional telephone lines. This approach has provided "connectivity" for many years, with a mail and message service to its 18,000 users when logged into WYLBUR, and it allowed these users to share programs and data. In the last two years the Computer Center has enhanced its ENTER MAIL facility and provided direct, integrated connection to the world-wide BITNET mail network.

In the world of computing, mail and file transfer on networking systems and other computer connections have also developed in other ways. The growth of LAN's created a large and growing number of Ethernet mail users at NIH, and the popularity of the TCP/IP communication protocols in the computer science community created a growing number of TCP/IP mail users throughout the academic world.

A major effort this year was put into interfacing 3Com 3Plus and TCP/IP mail systems to WYLBUR and BITNET mail in order to allow users on a 3Com 3Plus network to exchange mail with users on WYLBUR via ENTER MAIL, and to exchange mail with users at other BITNET nodes throughout the world. Users on a 3Com 3Plus network can also exchange mail with users on TCP/IP-based networks using the mainframe as a gateway; mail between users on different LAN's that use the same operating system is transmitted without going through the mainframe. Mail gateway software was tested throughout the spring and summer and has been integrated into WYLBUR, 3Com 3Plus, and TCP/IP. This facility also provides access to BITNET and the Internet, international data networks for mainframes (See figure 1).

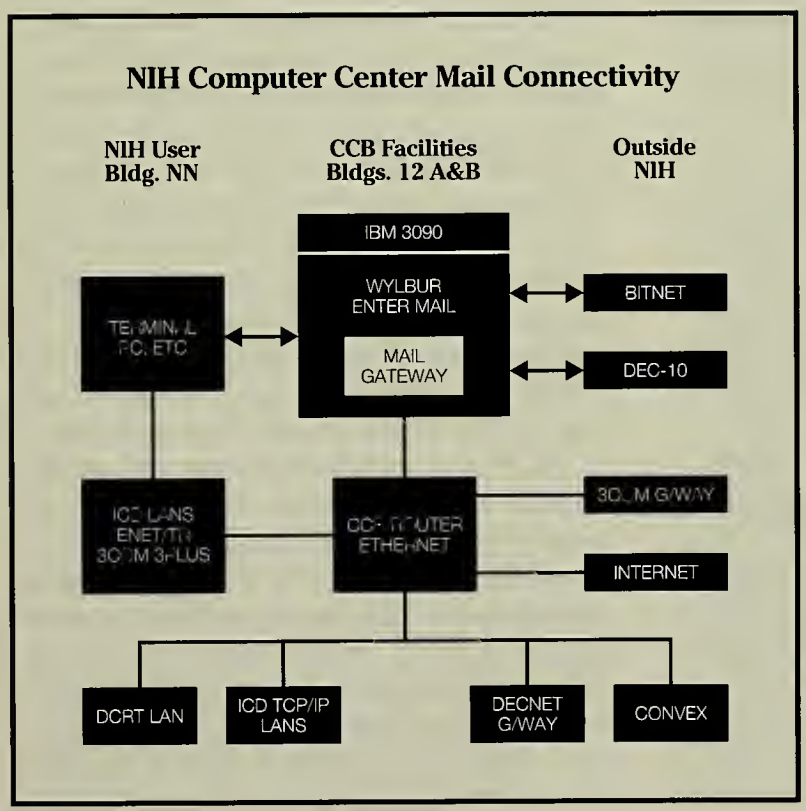


*Figure 1*

The connectivity plan enhanced access to BITNET and other national and international data networks connected to the computer utility. NIH use of BITNET has increased dramatically during the 3 years that the NIH Computer Center has been a member of the BITNET international digital communications network. BITNET was used about 3,500 times per month when it was first made available in the summer of 1986. Today, BITNET is accessed more than 70,000 times in a typical

month—a twentyfold increase.

Special interest LISTS proved useful to BITNET users at the NIH Computer utility during the past year. BITNET lists are electronic discussion groups catering to a variety of interests, ranging from the latest news in medical research to discussions on how to design data base systems. Over 130 lists are currently active. All LIST subscribers automatically receive, via electronic mail, copies of all comments and discussions contributed by other members of the group, and may submit their own contributions by sending electronic mail to the LIST.

New format conversion software was installed on the NIH Computer Utility in April to expand the capabilities of the BITNET connection by processing files using the NETDATA file transmission format. BITNET files transmitted in NETDATA format, to users at the NIH Computer Utility, are now automatically converted into a readable text format.

A DECnet-Internet gateway was installed in February to make it possible for users of the DECsystem-10 to access the Internet, a nationwide network of research institutions that communicate with each other using the TCP/IP family of protocols. A new product from Digital Equipment Corporation (DEC) provides a "gateway" for connections between the machines using DECnet protocols and machines that use TCP/IP protocols. Now, users of NIH computers that communicate with DECnet (e.g., DECsystem-10, VAX's, PC's, etc.) can communicate with systems on the Internet. Via a connection to the IBM System/370 they can also communicate and exchange files with the users of WYLBUR, 3Com networks, and BITNET systems.

Because users of the NIH Computer Utility often find it necessary to transfer files between a PC and the mainframe computers for processing or distribution via ENTER MAIL, the Computer Center took steps during FY89 to simplify this process. The Computer Center Branch developed a new release of MS-Kermit, the PC file transfer and terminal emulation software package supported by the NIH Computer Center. The new version, called Kermit 2.31+NIH1, was based on a collaborative effort by academic institutions coordinated by Columbia University and Utah State University and enhanced by Computer Center personnel to meet the unique needs of NIH computer users. Made available in March, the new version provides features that greatly simplify the process of using a personal computer to access mainframe services. It includes a simplified installation procedure that interacts with the user to create the necessary DOS files. Once Kermit has been installed, the user can access WYLBUR, TSO, or DB2 with one simple command; then all that the user has to type is a personal keyword. The terminal emulation facilities of Kermit are vastly improved in this version. Features include destructive backspace, fast scrolling, print suppress, and direct printing. The new facilities of Kermit 2.31+NIH1 allow users to easily create mail using a PC word processor, and send it through ENTER MAIL to BITNET or Internet addresses.

In addition, a new communications package for IBM PC's and compatible microcomputers, ProComm Plus, was given Level 3 support status on the NIH Computer Utility while the package is being evaluated for possible production service use. Developed by Datastorm Technologies, Inc. as a "shareware" product in the public domain, ProComm is a full-screen, menu-driven package with a built-in text editor, a scripting facility to automate communications procedures, a HELP

facility, and Kermit file transfer protocol capability.

The Computer Center's Evolving Technology (ET) unit continues to strive to identify, evaluate, recommend, and support new advanced connectivity products that leverage the strengths of both personal and mainframe computers for the scientific research and administrative support activities of NIH. The Computer Center has evaluated several other products that could improve the access to mainframe data and systems from personal computers. Products evaluated include: T-A-C Version 5.0 and PC/SQL Link, that allow PC spreadsheets and data bases to directly access mainframe data stored in the relational data base management system (DBMS), DB2; Execulink and Relay Gold, terminal emulators that incorporate advanced ease-of-use features; Attachmate EXTRA and the IBM 3270 Emulation Program, which allow groups of PC's on a local area network (LAN) to economically support high speed access to full-screen mainframe applications, such as the Administrative Data Base and DB2.

Communications between NIH and relevant external communities were further enhanced this year when PUBLIC user initials were made available on the NIH Computer Utility. PUBLIC initials allow appropriate individuals who are not registered users of the NIH Computer Utility to get valid access or submit information to data systems running on the utility. Thus, registered users of the NIH Computer Utility may now provide NIH relevant services to individuals who are not regular users of the utility.

This has proven to be an invaluable tool to scientists and administrators who must exchange information with diverse groups of people who are not users of the NIH Computer Utility. PUBLIC initials can be used to help eliminate the burden and delay of mailing guidelines and instructions or corrections; update pages to documents; and provide a means for non-NIH individuals to submit progress reports, mailing list updates, research papers, applications, and other documents, regardless of geographic location.

## The 3090 Transition

On September 28, 1988, the Computer Center signed a "Total System" contract with the International Business Machines Corporation (IBM) to provide continually evolving computational capability to NIH throughout the 1990's. The 10-year contract provides for hardware, software, and support services to the ever evolving computational needs of NIH. The whole processs of initiating, specifying, advertising, evaluating, and negotiating this contract spanned some three years. The contract called for a transition period during which the existing equipment was to be replaced and/or upgraded to new state-of-the-art technology without service interruption. This required careful planning and implementation.

The transition began within three weeks of contract award, when the first 3090 model 200 was upgraded to a model 300E in October 1988. The 300E includes three processors (there are two in a model 200) and each is approximately 10 percent faster than those in the model 200. The remaining processors were upgraded in November, December, January, March, and completed in April. The conversion of the shared disk storage facility, the second major phase of the transition, started immediately on completion of

the processor upgrades. Each new model 3380 device can store about 50 percent more data than the previous model E. The new disk modules are further enhanced by the use of state-of-the-art 3990-3 controllers. The new disks are physically similar to the older ones, containing 47,476 bytes per track and 15 tracks per cylinder, but they contain 2,655 cylinders per volume rather than 1,770 yielding the 50 percent increase in data storage, with no increase in floor space, power, or heat dissipation. The 3380K's also enhance overall system performance with faster seek times, while reducing the cost of online data storage. The DASD conversion required that over 350,000 user datasets stored on hundreds of volumes of FILE, CNTL, MSS, dedicated and TMP disks be transfered without affecting user access. The DASD upgrades started in May, continued with a new installation about every two weeks, and was completed when the eleventh module was installed in October.

Care and professionalism marked the processor transition installations. They went smoothly and on schedule. Thorough testing by many groups of the Computer Center and IBM ensured that the installations were completed without interruption of ongoing services.

As each processor was upgraded from a 3090-200 to a 3090-300E, other components of that computational kernel were also upgraded. Main memory was doubled from 64 million to 128 million bytes, and expanded storage was increased by a factor of 4, from 128 million bytes to 512 million bytes. The number of channels on each kernel was expanded from 48 to 64; each new channel having a transfer rate of 4.5 million bytes per second, 50 percent faster than the old channels. Each new kernel contains at least one vector facility providing high performance computation for scientific application. The "scratch space" disks used by the kernel were then upgraded to 3380 model K's.

Installation of the new 3990-3 disk controller with high speed cache memory made the NIH Computer Utility one of the first installations in the world to utilize these state-of-the-art controllers. The new controllers were first installed on the system development processor for a period of rigorous testing. After verifying their power and reliability, the new controllers were moved to a production system and became an integral part of all kernels and shared disk modules.

To provide compatibility with other computer installations that have not yet converted to the new 3480 cartridge tape drives, which are now standard at the NIH Computer Utility, eight new reel type tape drives were installed in February to replace aging units.

At the end of the transition the NIH Computer Utility consisted of 6 identical IBM 3090-300E processors in full production service, fully sharing 1.2 trillion bytes of central disk storage and over 40 gigabytes of disk "scratch space" dedicated to each processor.

## DECsystem-10 Transition

In May 1989, the Computer Center awarded a 5-year contract to Convex Computer Corporation to augment and eventually replace the venerable DECsystem-10 that is operated by the Laboratory Systems Unit of the Computer Center. The contract calls for Convex to provide a complete system based on a C220 processor. Throughout the 5-year life of the contract, provisions permit expansion of processing power, memory, disk and tape storage, and auxiliary equipment to

meet the expanding needs of NIH scientific computation.

Each Convex minisupercomputer central processor contains pipelined 64-bit scaler and vector units that can execute up to 8 instructions simultaneously. The system features 200-MB per second ports to memory and a 40-nanosecond cycle time. Physical memory can be expanded to 2 gigabytes to support large user applications, which may utilize the virtual memory limit of 4 gigabytes. The C220 is capable of expansion to three or four processors. Symmetrical parallel computing is provided via hardware-based "Automatic Self-Allocating Processor" (ASAP) technology. ASAP allows a single processor to request additional processors to execute portions of code that can run in parallel, thereby reducing the time needed to complete the job. FORTRAN and C compilers developed by Convex automatically optimize programs to use the vector and parallel processing capabilities of the computer architecture. Pascal, Lisp, and BASIC permit conventional scalar programming.

The new NIH Convex computer runs CONVEX UNIX, an operating system derived from Berkeley UNIX and enhanced to meet the rigorous demands of multiuser production computing environments.

Users of the DECsystem-10 accustomed to a wide variety of software development tools will find an even richer assortment under the UNIX operating system that runs on the Convex C220. NIH scientists who have made use of the GenBank database and the DNA sequence analysis programs on the DECsystem-10 will soon have one of the most highly regarded sequence analysis packages running on the new minisupercomputer. Even casual users will be able to make use of the high degree of connectivity and network access afforded by the UNIX operating system for mail, file transfer, and USENET bulletin boards.

Additional software on the Convex system implements the Digital Network Architecture protocols so that users of the DECsystem-10 and other campus DECnet nodes can transfer files to the new system, using standard DECnet file transfer commands.

Plans for the new CONVEX computer include a period of parallel operation of the DECsystem-10 and Convex to permit the orderly transfer of applications to the new Convex system. Written guidelines and personal assistance by the DECsystem-10 staff will ensure a smooth transition to the new system.

## Expanded System Resources and Capabilities for Users

The transition to new hardware under the IBM contract led to a beneficial increase in system resources and capabilities for users of the NIH Computer Utility.

The maximum size for data sets stored on CNTL, TMP, or MSS volumes was increased from 150 to 200 cylinders. This allows up to 142.4 million bytes to be stored in a single data set, an increase of one-third over the previous limit for CNTL and TMP data sets, and three times the previous maximum for MSS storage. The new limits facilitate the processing of large volumes of data and permit large files that were once restricted to magnetic tape to be moved to online storage. This allows tasks previously done as class B batch jobs to be done faster under job classes A or E, or interactively under TSO.

The limit on the number of tracks of scratch disk (SYSDA) space was increased from the previous limit of 30,000 tracks (approximately 1.4 gigabytes) to 100,000 tracks (approximately 4.75 gigabytes). This increase facilitates file processing, permits larger disk sorts, and eliminates many working storage constraints that have affected some applications in the past.

A new version of JES2 (Job Entry Subsystem) was installed on the system in January. Under the new version, the total number of jobs that a user can have in the system at once increased from 100 to 150, and the number of jobs that a user can have executing or awaiting execution increased from 40 to 50.

## New Software

In February, version 5.18 of SAS became the production version on the NIH Computer Utility. The new version solves a number of problems found in earlier versions and offers several enhancements and new functions. GAUSSIAN 86, a powerful computer tool for solving theoretical chemical problems, was made available on the NIH Computer Utility early in the year. Developed at Carnegie-Mellon University, this user-friendly program is one of the most widely used quantum chemistry programs in the field of computational chemistry.

FORTRAN version 2, release 3, was installed as the production version of FORTRAN early in the year. The release added several new features without requiring any changes in existing programs.

Release 10.0 of DFSORT replaced release 8.0 as the production version of the sort/merge program. The new release required no changes to the users' job control language and added new options while enhancing the efficiency of the program.

A new release of TELL-A-GRAF, a conversational computer graphics software package, was made available on the IBM system in February 1989. The enhanced and expanded new version supports new graphic devices and commands, performs faster and more efficiently, and is more cost-effective. Included in the new release is an option called TABLES that enables users to create complex tables with only a few simple commands.

CUECHART, an easy-to-use interactive program for making charts under TSO, was added to the NIH Computer Utility in February 1989. CUECHART provides an extensive library of over 1,100 predesigned charts and a system for prompting the user for required information and data, which is automatically converted to TELL-A-GRAF commands in order to plot the chart. CUECHART makes it easy to produce high quality charts without learning detailed TELL-A-GRAF commands.

New features and dramatic performance enhancements were made available when Version 2 Release 1 of DB2, the IBM relational data base system that has established itself as an industry standard, was installed on the NIH Computer Utility early in 1989. An important functional improvement was the addition of full referential integrity, a powerful design feature that maintains consistency among related data stored in different tables. Tests of the internal efficiency of the new version demonstrated that large sorts could be completed using only one-half the CPU time required under the old release.

## New Services

High quality, duplex (both sides of the paper) printed output on cut (individual sheet) paper is now available from the NIH Computer Utility as a result of newly installed, state-of-the-art IBM 3827 printers. The 3827 printer produces high quality output on 8 1/2 by 11 inch sheets in either vertical or horizontal orientation, and can handle both single- or double-sided printing. This not only eliminates ragged pin feed strips, but duplex printing saves money and simplifies the production and handling of large documents and large numbers of copies of smaller documents. Single-sided or duplex output can easily be produced by batch jobs. All character sets and densities available for 3800 printers can be used on the 3827 printers.

Users of the NIH Computer Utility now have interactive access to both individual and organizational sections of the NIH telephone directory and the NIH Computer Utility User Directory. This new facility is heavily used by grantees who must contract numerous grant administrators and research investigators at NIH relative to their grants or research. Anyone with a terminal or a PC and a modem can search, list, and download up-to-date information (name, address, telephone number, and organization) from either directory at no cost. Directory information is updated monthly, assuring that users get the most timely information available, and even those who are not registered users of the computer utility may access the online telephone directories by using a set of public initials established for this purpose. So enthusiastic was the initial response to the facility that the Computer Center subsequently expanded it to include the organizational pages of the NIH telephone directory.

## Rate Reductions and Refunds

In keeping with its 21-year tradition, the NIH Computer Center again offered numerous rate reductions to users of the NIH Computer Utility during the past fiscal year. Some of these reductions were brought about by the installation of more efficient hardware made available under the new Total System Contract.

The first set of rate reductions for both IBM System 370 and DECsystem-10 users was put into effect on October 1, 1988. These reductions included the following:

- Elimination of the I/O component of the charge for the TSO interactive system, and an 8 percent cut in the charge for CPU time, from $1.25 to $1.15 per CPU second.

- Reduction in the cost of a machine unit for batch processing to $.27 for all classes of batch processing services.

- Lowering of the point at which the REGION charge rate is reduced, to 2.5 MB and elimination of additional charges for all region above 6 MB, facilitating the assault on very large scientific computational problems.

These changes reduced user costs by a total of nearly $1 million per year.

- Reduction of charges for timesharing and batch services on the DECsystem-10 by 10 percent, from $.55 to $.50 per machine unit. The second set of rate reductions was implemented on March 1, 1989, when some rates were reduced even further and rates for additional services were reduced. These reductions included the following:

- Another reduction in the charge for TSO CPU time, bringing the rate down to $1.05 per second, a reduction of almost 9 percent.

- Reduction in rates for WYLBUR from $1.15 per second to $1.05 per second.

- Lowering the rate for storing data on public disk space from $.014 per track day to $.013 per track day, a decrease of about 7 percent. In addition, charges for dedicated disk volumes were changed to a per-volume basis rather than a per-track-used basis, thus reducing the dedicated volume charge for most users.

- A rate reduction to users of the IMS data base system of almost 12 percent for IMS services, when the IMS transaction charge was reduced from $.17 per transaction to $.15 per transaction.

- Elimination of the penalty factor for G and H class scientific jobs on the IBM vector facility that do not meet minimum vector usage standards. This relaxation of eligibility requirements for the lower rates means that more scientific computing jobs can now take advantage of the processing power available through the Vector Facility.

- A reduction in lease rates for CRT and hardcopy interactive terminals by an average of 16.6 percent. Actual savings ranged from 7 percent for an impact printing terminal to 38 percent for a letter quality printing terminal.

The cost effectiveness of the new state-of-the-art equipment and software, the economy of scale, effective management policies, and innovative users, all contributed to a cost saving of $10 million during the first half of the year. This saving was shared by all users of both the IBM 370 and the DECsystem-10 in May when it was rebated to user accounts. Factors such as these enable the NIH Computer Center to provide the most advanced computing services available anywhere at the lowest possible cost.

## Documentation, Training, and Computer Assistance

Online documentation was introduced in October 1988 to provide an easier way for users to access technical documentation. With the command ENTER DOC, users can now select a document, look up items in its table of contents or index, and view the applicable portions of the text at a terminal or personal computer. ENTER DOC provided access to the Computer Center "Users Guide," the "DB2 User's Guide," "Using Tapes at NIH," and three documents from the Personal Computing Branch: the "PWO Printer Guide," "Column Manipulation in WordPerfect," and "Retrieving Documents in WordPerfect." A prototype system, ENTER DOC is the Computer Center's first step in providing online documentation.

During the year the Computer Center "Users Guide" was updated twice, 17 revised technical documents were published, and 53 other updates were distributed to users. The Automated Documentation Service was used by 7,097 users, and 137,518 copies of technical publications were sent to those users and to individual requesters. Ordering publications online was popular with users; ENTER PUBWARE was used 4,308 times to order documentation. INTERFACE, the Computer Center's technical newsletter, published seven issues, including the annual index.

Creative training opportunities continued to be an important function of the NIH Computer Center Training Unit. This year 39 different courses were offered in the formal classroom training program. Of the students who requested training, 1,075, (or 82 percent) were accepted into the formal training program. Informal seminars on topics such as BITNET and Kermit 2.31+NIH1 were attended by more than 740 people.

Independent study courses continued to be popular among users who could not take the formal training courses. To meet growing training needs, a new Assisted By Computer (ABC) course, "Introduction to Partitioned Data Sets," was added. This brings the total number of ABC courses offered to eight. More than 3,400 people took ABC courses during the year.

To better meet the needs of users seeking assistance, the Computer Center began offering consulting appointments, an opportunity for users to discuss specific requirements or difficulties with members of the professional staff of the Computer Center on a one-to-one basis. Over 200 people took advantage of this service, receiving assistance in such diverse areas as uploading files from a PC, using BITNET or ENTER MAIL, designing DB2 applications, creating VSAM data sets, writing WYLBUR command procedures, and designing program logic for research projects.

This year, Computer Center consultants handled 12,079 requests for customer assistance and researched and answered over 3,149 programmer trouble reports. Changes to improve the performance and reliability of the operating system required the implementation of 66 SYSGENS (software reconfigurations). Approximately 22,500 fixes, both preventive and corrective, were tested and applied to the system and 48 new releases of current software packages were installed.

## Future Outlook

The coming year will have two major focal points: introduction and expansion of support for UNIX (Trademark, AT&T) and an emphasis on campus-wide connectivity enhancements.

The introduction and expansion of support for UNIX mirrors the rapidly growing acceptance of UNIX as the de facto environment of choice for scientific computing. UNIX provides an easy-to-use, platform-free environment that facilitates development of scientific applications and easy movement of applications and data between hardware platforms.

Communications between and among the various scientific workstations, mainframes and supercomputers running UNIX are simple and standardized and will greatly assist collaborative communication and research.

The UNIX-based Convex minisupercomputer will assume the workload of, and replace, the DECsystem-10, and establish one base on which to provide even further improvements in scientific computing capability. The Convex system supports DECnet networking capability to facilitate conversion of existing DECsystem-10 applications and communication with existing DECnet networks (e.g., VAX-based machines) as well as the full TCP/IP suite of communications protocols. The Convex system provides both FORTRAN and C compilers capable of using the high-performance vector and parallel capabilities of the multi-processor Convex system, and a rich base of scientific applications packages.

The introduction of UNIX in the IBM environment will continue under study as AIX/370, a UNIX operating system for the IBM System 3090, will be evaluated for installation on one of the 3090 multiprocessor systems. The 3090 is a high-performance system with both vector and parallel computing capabilities that, when coupled with over 1 trillion bytes of online storage, can address the large-scale computing needs of NIH research. AIX/370 was designed to be compatible with both the System V version of UNIX from AT&T, and version 4.3 of Berkeley UNIX. The AIX/370 system provides FORTRAN and C compilers compatible with the Convex system and capable of using the full capability of the 3090, full connectivity via the TCP/IP protocol suite, access to the large online storage environment of the IBM systems, and a potential for a large base of scientific applications packages. The AIX system is being evaluated by the Computer Center in conjunction with a number of NIH researchers with requirements for extensive amounts of computational power.

The combination of the Convex and IBM systems provides a multiple, supercomputer capability for addressing virtually any computational and data storage need in the NIH research environment.

Connectivity will continue to be a major area of effort as the existing connectivity facilities are expanded into one of the most comprehensive personal computer, scientific workstation and local area network (LAN) systems to be found. The planned connectivity facilities will permit users to exchange mail and files with other mainframes, PC's, laboratory workstations, and LAN's connected to the utility, even if the exchange is between different operating systems or equipment made by different manufacturers. When the connectivity plan is fully implemented, users with personal computers or workstations will be able to access interactive services and batch processing facilities over a high-performance network. The network will use a variety of media, including twisted-pair wire, coaxial cable, fiber cable, T1 and T2 service, and will connect to the NIH Campus Area Network when it is available.

As the Computer Center moves into the future, it will continue to develop new ways to provide the most modern computer facilities available at the lowest possible price. The Computer Center staff remains firmly committed to that goal.

# Personal Computing Branch

*David C. Songco, Chief*

The Personal Computing Branch (PCB) provides guidance and support to scientists and administrators throughout NIH in the effective use of personal computing technology, including workstations, networks, and associated technology. PCB works closely with other DCRT labs and branches and with NIH groups outside DCRT to develop and maintain a broad and multifaceted program to meet the needs of the NIH user community.

The PCB staff of 22 comprises computer specialists, engineers, and instructors. It is organized into four sections, each focusing on a specific area of technology. Robert Romanoff heads the Administrative Technology Section, which currently tracks the IBM PC and compatible family of computers. Brian McLaughlin leads the user training program and also heads the Scientific Technology Section. This section focuses on support and guidance for biomedical research programs at NIH. It also forms the basis for supporting the rapidly growing Apple Macintosh community at NIH. The Communications Technology Section, lead by James Del Priore, specializes in the rapidly changing and complex local area network technology. Bonita Condon heads the Consulting Services Section, which provides guidance and support for PC-based word processing and database applications.

## The NIH User Resource Center Helps Users Help Themselves

PCB sponsors the NIH User Resource Center (URC) in collaboration with the Division of Personnel Management and the Division of Management Policy. Clark Collins represents PCB in the URC on a daily basis. The URC is a multipurpose center that contains both PC and Macintosh workstations equipped with a variety of laser and matrix printers as well as a full range of software packages supported by PCB. During FY89, new workstations featuring PCB-supported IBM Micro Channel Architecture (MCA) were made available in the URC. Additional Macintosh machines and resources were also put in the URC to help keep up with heavy demand by the user community. The URC also has self-study courses available on personal computer and videotape recorder. More than 3,000 NIH employees took self-study courses or used the personal computer equipment in the URC in FY89.

In FY89, PCB helped the URC plan, coordinate, and organize the NIH User Resource Center PC Fair 88. The fair, held in November, gave NIH employees an opportunity to share their PC experiences with one another. The fair planning committee recruited over 100 NIH employees who either assisted with the logistics of the fair or demonstrated PC applications. More than 1,000 NIH employees attended the fair.

## Lead User Program Is Revamped in FY89

The centralized support provided by PCB is enhanced by the PC lead user program. This program consists of more than 200 BID-nominated NIH employees who assist PCB by serving as the first line of support within the NIH personal computer user community. Lead users answer PC-related questions, provide training, and disseminate PC-related information on behalf of PCB. Lead users are also encouraged to participate in the PCB training program as associate instructors.

This year, PCB staff, under the leadership of Nancy Walther and Lori Collins, worked closely with lead users to improve the lead user program. A profile of individual lead users was developed so that inactive lead users could be withdrawn from the program and certain lead users could be certified as qualified in particular areas. Certification is part of a renewed effort to encourage users to seek help from lead users at the local level and to ensure the help there is adequate. Our renewed emphasis on local support is driven by the rapidly increasing demand for support and guidance as the size of the user community grows and the number and complexity of microcomputer products increases.

PCB monthly "Lead User Topic Sessions" continued throughout FY89. The meetings, cochaired by a PCB staff member and a BID representative, consist of demonstrations, tutorials, and discussions of computing topics likely to help lead users better serve members of their BID's. Topics presented and discussed at the FY89 meetings included portable computers, database management software, graphics programs, forms management, communications, spreadsheet applications, and aspects of personal computer architecture such as memory management.

## Quality Training Promotes Effective Use of Computer Technology

The PCB-coordinated microcomputer training program continued to expand and improve in FY89. A total of 2,600 students enrolled in 184 sessions of 34 different courses. Vendors selected by PCB and the NIH Training Center presented 14 of the courses. The remaining 20 courses were taught by PCB, DCRT, or other NIH staff. Roughly half the PCB/DCRT courses were administered by the DCRT Training Unit and offered without fee. There was a charge of $50 per day for courses administered by the NIH Training Center and taught by DCRT staff. The cooperative arrangement between the NIH Training Center and DCRT made possible the addition of several PCB courses and classrooms and made manageable the substantial administrative overhead associated with running a diverse and extensive training program.

Two new network courses were introduced in FY89: a one-day "Introduction to PC Networking" class and a three-day "Network Administrator" class. Both courses were developed by a contractor under the direction of PCB staff. Macintosh training was refined and greatly expanded in FY89. Introductory classes were consolidated into a popular one-day "Welcome to Macintosh" class, offered twice weekly. Vendors have been hired by DPM to teach application-level courses for PCB-supported Macintosh software packages.

A vital factor in the high quality of PCB training is the PCB associate instructor program. This program illustrates the PCB philosophy of maintaining a close working relationship with the user community by involving it in all aspects of support and guidance. In FY89, more than 90 lead users volunteered to serve as associate instructors for an average of five days during the year, making it possible to have both a main instructor and two assistants for each hands-on class.

## Advances in Technology

During FY89, technological advances in PCB-supported PS/2 personal computer systems based on the IBM Micro Channel Architecture continued to bring increases in performance together with reductions in cost. The PCB evaluated and announced support for several new adapter boards designed to work with MCA. Many new software applications, as well as updated versions of currently supported software products, were likewise evaluated by PCB staff, lead users, and other specialists within DCRT. Early testing of Operating System/2 (OS/2) was continued in an effort to determine its potential at NIH. OS/2 versions of some currently supported software packages were released in late FY89, but a full complement of products will not be available until at least the third quarter of FY90. Current and forthcoming OS/2 applications offer the promise of both increased functionality and a more efficient user interface, leading to greater productivity at the workstation. PCB and other components of DCRT will be actively involved in selecting the best mix of OS/2 products to provide a suitable computing base from which NIH users can operate.

NIH interest in presentation graphics and desktop publishing remained high in FY89. In response to technological advances in some peripheral devices such as printers and scanners, PCB evaluated and announced support for several new products that offer greater speed and higher resolution at reduced cost. Many PC-based products for the design, generation, and filling out of forms emerged in FY89. Widespread interest in this technology at NIH led to a cooperative effort between PCB, other components of DCRT, and the Division of Management Policy to evaluate NIH's forms requirements. The effort is expected to result in the development of a set of guidelines and recommended products to provide NIH-wide consistency in electronic forms generation.

## Macintosh Support Expands

PCB estimates the number of NIH Macintosh computer users to have doubled during the first year of PCB support. In FY89, PCB Macintosh support expanded to keep pace with the growing needs of the NIH Macintosh community. Much of the support effort went into evaluating new Macintosh products, writing articles for the PCB's technical publication, PCBriefs, and providing telephone hotline support. The staff now averages approximately 100 Macintosh consults a month. The branch published an initial list of PCB-supported Macintosh products to help provide guidance in product selection. That was followed by publication of the PCB Macintosh Product Information Guide, with detailed information about PCB-supported Macintosh products and NIH resources. Through an active conference on PCB's electronic bulletin board service—PCBull—Macintosh users were given access to timely Macintosh-related information and much free software.

PCB also continued to give full support to the Biomedical Research Macintosh User's Group (BRMUG), a coalition of NIH Macintosh users. Mike Basham of PCB quickly emerged as one of the BRMUG leaders and did an outstanding job coordinating activities in FY89. Through monthly meetings, special events, product demos, and a public-domain software disk, BRMUG provides a lively forum for the exchange of Macintosh information at NIH.

## Local Network Advances

This year PCB continued to play an active role in the development of locally managed networks at NIH. Judy Fabrikant provided technical leadership to local server and network administrators within DCRT, and staff met monthly with representatives of locally managed networks to exchange information and to coordinate both central and local support services. Day-to-day support and guidance was provided via electronic communications established among the networks.

## Connectivity Continues to Challenge PCB

PCB provides active leadership in NIH connectivity activities via the DCRT connectivity group. This group of DCRT technical staff was established in FY87 to help ensure a coordinated effort among the many DCRT projects underway involving communication between the different levels of computing at NIH. The group publishes reports that contain descriptions of DCRT policies, services, and technical guidance designed to assist NIH specialists, managers, and planners involved in the difficult task of implementing effective communication between workstations and other computing devices. In FY89, a "Network Naming Conventions Recommendation" was published to promote effective and efficient communications among the diverse networks at NIH. The TCP/IP, XNS, and DecNet protocols were covered in the document. The networking user community at NIH participated in the generation and review of these conventions.

## Dissemination of Computer-Related Information

The results of product tests and evaluations as well as other helpful tips derived from the user consulting process are made available to users in a variety of mediums. The value of this information can be increased substantially if it is disseminated to the NIH community in a timely manner. With this goal in mind, PCB staff made several changes to the process of reporting computer-related information to users.

To get information to the broadest spectrum of users as quickly as possible, PCB, under the direction of publications coordinator Dan Zoll, set up an electronic bulletin board system that was christened PCBull. NIH employees can access PCBull from their own workstations via telephone communications. The system contains descriptions of all PCB-supported products and has a large variety of other informational files such as product updates and usage tips. In addition, the PCB Utility Files and other public-domain programs for IBM PC's and Apple Macintoshes are available on the system for downloading. PCBull also allows users to ask questions of PCB online. PCBull was an instant success and has already become an integral part of our support program.

At the same time, PCB expanded its use of electronic communication with network users at NIH. Since these users, by definition, have access to high-speed electronic communication, the PCB made use of direct internetwork communications to provide support for this group. In this manner, PCB staff specialists on the DCRT LAN accept questions from and provide guidance to their BID counterparts on locally managed networks. This internetwork communication resulted in the timely distribution of con-

To support the auditing efforts of the Clinical Center Medical Records Department (MRD), DMB has converted its data collection procedures to an online LAN environment. This system collects and validates all MRD data and forwards it to the mainframe for storage and reporting. The next step in this process is to convert the mainframe database to the LAN and establish the feasibility of running the entire system in a shared PC environment.

The Clinical Information Utility (CIU) was developed during the 1970's as an historical archive of clinical information for research. CIU gathers its data from the Medical Information System (MIS), the MRD and the various service organizations in the Clinical Center (CC). Over the years, multimillions of records have been archived and are made available for search and display with overnight turnaround of most requests. The Information Systems Department, CC, monitors and authorizes all users of CIU data. The CIU automatically tracks and reports each access of the data base. During FY89, DMB handled approximately 2,300 requests for information from the CIU.

Enhancements to the CIU this year included addition of historical surgical pathology data and expansion in the number and size of ICDA codes.

During FY89, DMB began to provide a limited personal computing service that supports analysis of nucleic acid and protein sequences. This service has included (1) running data base searches, (2) advising investigators regarding different approaches for sequence analysis, and (3) using a variety of commercial, public domain, and tailor-made application programs to fill their needs. Both the GenBank and the Protein Information Resource (PIR) are available in original format and DNASTAR format. Homology searches are supported by way of the Lipman/Pearson FASTA/TFASTA programs. Sequence related work is supported by the MBUG programs and the DNASTAR sequence analysis package developed at University of Wisconsin. This service has already been used by several institutes and is being offered NIH-wide on a "time available" basis.

To track Freedom of Information (FOI) Act requests, responses, hours, and costs, DMB developed a relational data base system for the FOI office, OD. This new facility replaced the old tracking system and provides the FOI office with online query capability.

Last year, DMB implemented a management information system for the NICHD. The system supports the management of NICHD extramural research projects. Data for the Child Health Information Portfolio System (CHIPS) is available online and consists of DRG's IMPAC system data as well as institute data. To support the increasing use of personal computers in NICHD, a facility was developed that permits CHIPS data to be downloaded and processed with various PC software packages. Training of NICHD personnel in the use of CHIPS is currently underway. Additional enchancements continue on CHIPS to provide better support for the grants management activities.

DMB provided additional support on more than 30 other projects. To support NIMH studies of seasonal affective disorder, DMB assisted in the analysis of the timing and sequence of symptoms during onset. For the OD, NIH, DMB developed a full time equivalency system that brings together hours, salaries, and fringe benefits from the personnel and financial management systems. The branch provided continuing support

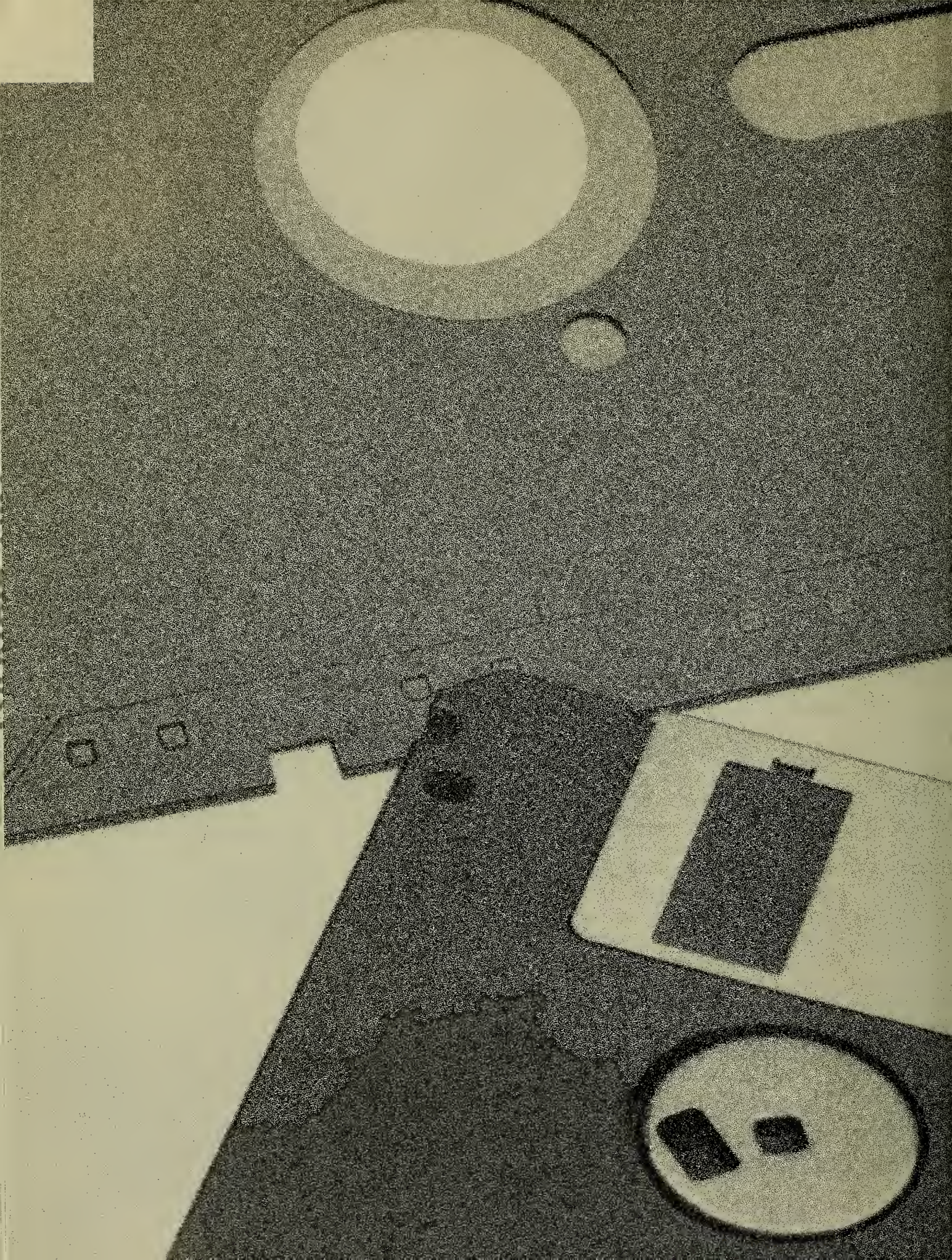
for (1) the nutrition grants monitoring system, (2) the visiting scientist program, (3) the cardiology/cardiac surgery tracking system, (4) the NCI seroepidemiology analysis system as well as its grant literature system and (5) Fogarty scholars assistance. The branch also provided final documentation and contractor training when it turned over the NIAID database system for HIV vaccine trials. A new system was implemented for pulmonary care and analysis to permit online entry of multiple sets of data for a given patient on the same day.

## Future Plans

During the coming fiscal year, DMB plans to have an ADB maintenance contract in place and begin training contractor personnel on ADB systems. As soon as branch personnel can be released from maintenance tasks, the ADB development effort will be finalized using the current Information Management System (IMS). Staff will then be reassigned to concentrate on management information processing and evaluation of hardware and software to replace the IMS data base management system. The evaluation effort will probably not begin until mid-FY91 and the initial studies will concentrate on selecting candidate solutions for evaluation.

In order to synchronize the efforts of the Personal Computing Branch (PCB) and DMB, the two branches will work closely together during the next several months to (1) define the types of support each will offer to the personal computer user, (2) establish procedures for evaluating hardware and software and for supporting and applying highly useful products, and (3) most importantly, establish standard approaches for keeping each other informed as changes occur and major new facilities become available. The compelling reason for these synchronization efforts is that the NIH community deserves to know the types of personal computing support that will be offered centrally and the manner in which that support will be provided.

The increasing pace that new technology becomes available presents a continuing challenge. One must be prepared to develop applications using proven methods with a history of success and known reliability, but in a manner that assumes a level of upward compatibility. This requires a proper balance between the desire to implement new technologies and the realities of adequate support and practical application. It is the DMB plan to assure early evaluation of new technologies, train its technical staff on preferred solutions and fit the best mix of available hardware and software in an effort to solve computational problems.



# Laboratory of Applied Studies

*John E. Fletcher, Ph.D., Chief*

The Laboratory of Applied Studies (LAS) is a multidisciplinary laboratory that includes mathematicians, computer scientists, physicians, and engineers. LAS operates in a task-oriented mode rather than in areas of separate disciplines. The Laboratory's approach to research problems is to examine underlying scientific principles, identify appropriate mathematical and engineering concepts, and use computing systems to carry out research objectives. Many projects in this laboratory are collaborative efforts with bench scientists and clinical investigators at NIH or at other research centers.

The scope of LAS investigations ranges from direct involvement in clinical and laboratory activities to the abstract development of mathematical methods and computer algorithms essential to quantitative computer modeling. The resulting software is made available as a general research tool. These activities are carried out administratively by two sections, the Medical Applications Section, which includes physician-scientists, electronics engineers, and computer systems analysts; and the Applied Mathematics Section, which includes specialists in applied mathematics, computer science, and computer modeling. Guest workers and university students or other guest volunteers are also involved in LAS projects.

## Collaboration Progress

In FY89 LAS continued the trend to enlarge the computing function of desktop computing systems to connect with the mainframe and to other facilities via the local area network. Many older obsolete systems such as the DeAnza image analysis facility and the LSI-11 signal processing facility are being replaced by newer systems that reflect modern technology. As in FY88 the shortage of personnel slowed the pace of most research projects, and some, such as the clinical collaborations through the medical staff fellowship program, were moved out of NIH to investigators elsewhere.

Most of the text preparation and scientific manuscripts have been converted from a WYLBUR base to PC Wordstar or PC-based scientific text editors such as EXP or to a PC-based VMAP. Manuscript turnaround and editing time has been reduced by the acquisition of laser printers and access to the DCRT LAN. The laboratory is continuing to expand its PC systems to take advantage of improvements in technology and software as they become available.

The **Medical Applications Section** developed a statistical model for the diagnostic power of ECG algorithms that was tested using the Framingham Heart Study results. The staff also designed new filters for baseline wander and made comparisons using older standard techniques. The new methods were demonstratively superior.

The laboratory implemented a grammar for the molecular biology database command language interface. It is currently undergoing testing with the AIDS sequence database. The staff is reviewing search techniques for incorporation into the system interface.

The signal processing activities are being modernized and centralized in a MAC II system to facilitate use and remove dependence on obsolete systems. This facility will be completed in FY90.

The image processing project on spinal cord injury with Children's Hospital was completed in

FY89. Collaborative activities with the Clinical Center Nuclear Medicine Department featured acquisition of systems to collect three-dimensional image data. Work is underway to use these systems for three-dimensional visualization, multimodality registration, and to explore interactive algorithms for analysis of large volumes of data.

The **Applied Mathematics Section** carried out a number of new applications of both least squares and SVD model/data fitting. Most involve laboratory or in vitro systems, but some may be extended to noninvasive diagnostic or clinical applications. Several analytical efforts were similarly productive resulting in publications and/or invited presentations.

The differential equations software development slowed considerably in FY89 due, in part, to the serious illness of the principal investigator, and the loss of a position for a part-time employee in support of this effort. The staff is writing software for PDE solution for the newer PS-2 systems and working versions of the FORTRAN software are functioning. Limited efforts are being made to improve the user-friendliness of the systems and to generalize its applicability to a wider class of problems.

J. Fletcher continued to chair the Mathematics and Computer Science Department of the Foundation for Advanced Education in the Sciences (FAES), and was active as a consultant on mathematical methods and software for ordinary and partial differential equations. He also served on the National Council of Teachers of Mathematics, commission for mathematics curriculum standards in FY89.

R. Shrager continued as consultant to NIH biochemists in areas of data analysis and experimental design, particularly in areas involving model fitting and digital filtering.

J. Bailey continues as consultant on Common Standards for Quantitative Electrocardiography, a program in medicine and public health sponsored by the European Economic Community. He is Secretary-Treasurer of the International Society for Computerized Electrocardiology, and has been appointed to the ECG Committee of the Association for the Advancement of Medical Instrumentation and serves as chairman of its Subcommittee on Ambulatory Electrocardiography.

Margaret Douglas chairs the Women's Advisory Committee and in that capacity managed the logistics for Career Day; performed analysis of workforce data and wrote the committee's portion of the annual Affirmative Employment Report; organized NIH Women's History Month observance; and served on the selection review board to evaluate candidates for the Federal Women's Program manager position.

## Future Tools, Trends

The applied mathematics activities will seek software for functional display in two and three dimensions. The laboratory will emphasize improving model specification to analysis software and the detection of ill-posed problems prior to attempted solution. An exploration of both UNIX-based and DOS-based versions of MLAB may begin in FY90. With the phaseout of the present DECsystem-10, many of the present modeling and analysis techniques will be reexamined in the context of the DECsystem-10 replacement.

The nonlinear modeling project will consider Raman temperature monitoring applied to

systems more complex than heavy water. Modeling efforts in progress will continue at whatever pace progress allows.

Simulations with the immune system model will examine the characteristics of targeted toxins or cell-specific killers as anti-AIDS interventions. Development of the global model will continue as new concepts can be translated into mathematical form and necessary parameter values can be estimated.

The signal processing laboratory will be completely reconfigured in FY90 and integrated with the image processing facilities. The image processing activities will concentrate on the design of an image transmission network and standards for image processing workstations.

In the molecular database project, the staff will implement the search algorithm(s) that are found most appropriate into the database, and add the calling sequences for analysis programs.

## Project Reports

### Mathematical Models of Binding Equilibria

*John E. Fletcher, Ph.D.*
*with E.W. Richards, Ph.D. (Baptist Research Centers, Birmingham, Alabama)*

The staff is conducting a study of mathematical models of ligand-receptor and ligand-macromolecule binding at equilibrium. Appropriateness of various model fitting criteria are studied and general guidelines and computational algorithms are designed for computer-aided interactive model fitting.

During FY88, a detailed experimental study was made of the binding of 16-[9-anthroyloxy] palmitoyl-CoA to BSA. During the data collection phase, the principal investigator accepted a new position at The Baptist Medical Research Center in Birmingham, Alabama. In FY89, the investigators accomplished analysis of the binding data, and prepared manuscripts describing the analysis of binding studies with BSA that are undergoing minor revision before final publication. It is uncertain whether additional studies will be attempted as a follow up to this effort.

### The Solution of Reaction Diffusion Systems in Biology

*John E. Fletcher, Ph.D.*
*with J. Weinstein (NCI), G.H. Weiss (DCRT/PSL)*

Investigators are developing numerical methods and mathematical software to solve ordinary and partial differential equations that describe dynamic physiological processes.

During FY89, software to solve single and coupled systems of reaction-diffusion equations was revised and installed on the PS/2 Model 60 using Ryan MacFarland FORTRAN. The PDECOL system is also installed on the same machine. The system is now being expanded to a wider class of problems, and its operating characteristics are being explored. The lack of user-friendly two- and three-dimensinal graphics for data display is disadvantage to PC use of this software.

The monoclonal antibody application (NCI) has been expanded to cylindrical and spherical geometries and to consider a bivalent antibody. Flow transport has also been added to the model's features. Manuscripts detailing the findings have been submitted for publication. Future studies will examine the geometric effects of tumor shape and density and lymphatic or other tissue flows on antibody activity. Other applications in progress are modeling of biased reptation for gel electrophoresis of DNA to determine optimum separation conditions.

**Cellular Kinetics Models of the Human Immune System (An Investigation of HIV-like Infections in a Model Immune System and its Response to Opportunistic Pathogens)**

John E. Fletcher, Ph.D.
with J.J. Bailey, M.D.; R.I. Shrager, M.S. (DCRT/LAS); W.L. Jackson, M.D., Ph.D. (NHLBI/CIPCB)

LAS is attempting to model the human immune system and the kinetics of its many complex interacting components, i.e., precursors, CD4+ T-cells, B-cells, T-cytotoxic, T-suppressors, killer monocytes, interleukins, or lymphokines, by means of a computer model, i.e., a system of nonlinearly coupled ordinary differential equations. An appropriately constructed and validated model should lead to experiments and interventions that can guide the use of treatments and vaccines, and promote the understanding of how the immune system might be manipulated to increase its effectiveness in preventing or neutralizing pathogenic infections.

In FY89, the identification of parameter values has been the primary function of this activity. The staff has made presentations describing the preliminary versions of the model and carried out some parametric studies that illustrate the model's behavior. It is expected that the investigators will refine the parameters and parametric studies as understanding of the various cell kinetic interactions improve. The staff is seeking closer collaboration with laboratory and clinical investigators to aid in the refinement process.

The staff has also carried out simulation studies of in vitro AIDS experiments with CD4+ T-cells and made presentations describing the modeling efforts. A guest worker has reviewed the literature on global models of the human immune functions and explored a preliminary version of a global model.

In FY90, researchers will explore the expansion of the in vitro simulation model to include syncitial formation and to include CD4 attractant toxins as a cytotoxic killer of infected cells. As in the global models, reasonable parameter ranges for the kinetic interaction parameters are difficult to obtain.

A presentation at a national meeting and an invited presentation at an international meeting on AIDS and theoretical immunology were well received. Manuscripts detailing these inquiries have been submitted for publication.

**Mathematical and Computational Methods for Solving Nonlinear Equations**

*Richard I. Shrager*
*with P. McPhie, Ph.D. (NIDDK/LBM); R. Berger, Ph.D. (NHLBI); N. Lewis, Ph.D. (NIDDK); I. Levin, Ph.D. (NIDDK); R. Hendler, Ph.D. (NHLBI/LC); R. Winslow, M.D.; M. Marini, Ph.D. (LAIR Blood Res.); M. Leon, M.D. (NHLBI); M. Gitterman, Ph.D. (Bar-Ilan University); G.H. Weiss, Ph.D. (DCRT/PSL); G. Campbell, Ph.D. (DCRT/LSM)*

This project is providing NIH investigators with mathematical tools for insight, analysis, and solution of complex equations that arise in the modeling of biological systems. To facilitate these efforts, LAS develops mathematical methods that are accessible to investigators from many disciplines. Software packages that result from these developments are made available to the research community as general research tools.

**FY89 Progress**

*with* Peter McPhie (NIDDK). The staff has submitted a paper on the thermal unfolding of swine fibrinogen. Models that best fit the circular dichroism data strongly suggest that unfolding occurs independently in two distinct regions of the molecule.

*with* R. Berger (NHLBI). Infrared (IR) spectroscopy, as analyzed by partial least squares, is being investigated as a probe of the ATP-ADP-PO4 system. If this IR technique can be refined, it will have many other uses, including noninvasive clinical applications.

*with* N. Lewis and I. Levin (NIDDK). Using Raman spectra of heavy water, as analyzed by singular value decomposition (SVD), the temperature of a solution was monitored. Raman spectroscopy is complementary to IR spectroscopy in that signals weak in one will often be strong in the other. The staff made a presentation on the preliminary aspects of this work in February.

*with* R. Hendler (NHLBI). A project studying cytochromes has produced two new mathematical modeling efforts. One accounts for the internal and external binding of tetraphenyl phosphonium in lipid vesicles and eventually in mitochondria. Another deduces voltages and redox states of cytochrome oxidase in coulemetric experiments. The staff is developing equilibrium binding models for both systems.

*with* R. Winslow and M. Marini (Lair Blood Res.) The Adair model of hemoglobin-oxygen binding is producing physically unrealistic results. Alternative models are now being considered that should clarify how the configuration of the hemoglobin molecule changes.

*with* M. Leon (NHLBI). A realtime least-squares procedure is being considered for use with fluorescence-guided laser angioplasty as an aiming technique to decide when the laser should be fired.

*with* M. Gitterman (Bar-Ilan University) and G.H. Weiss (DCRT/PSL). Researchers are investigating the influence of noise on systems described by the Mathieu equation, which is used in many scientific applications including superimposed biological rhythms.

*with* G. Campbell (DCRT/LSM). A paper surveying the robustness of weighted estimators was published. This work, which analyzed simulated data mainly by linear programming, is intended to guide the biochemist toward more reliable robust curve-fitting methods.

An abstract of an invited talk has been accepted on the estimation of weights in singular value decomposition (SVD).

Analytical work in this area is intended to improve the reliability of parameters deduced by SVD in biochemical analysis. SVD is used as an alternative to direct matrix least-squares or other methods for resolving components from multi-spectra measurements.

**Future Tools, Trends**

Raman temperature monitoring will be applied to systems more complex than heavy water. Animal experiments designed to validate the method of spectral edge frequency monitoring in anesthesia are scheduled to begin. The staff is developing a paper on the Mathieu equation project. Modeling of hemoglobin and cytochrome systems will continue to explore the alternative concepts. MLAB consultations will begin as soon as the new systems have been accepted, and the staff will be developing courses following acceptance testing of new software and hardware.

**Computer-Aided Analysis of Electrocardiography**

*James J. Bailey, M.D.*
*with M.R. Horton, M.Sc. (DCRT/LAS); R.I. Shrager (DCRT/LAS); G. Campbell (DCRT/LAS); Levy D. (Framingham Heart Study )(NHLBI); Haisty, K.W. Jr., (Cardiology Department, Bowman-Gray School of Medicine); Willems, J.L. (University Hospital St. Rafael, CSE, Leuven, Belgium)*

These studies evaluate the prognostic power of the electrocardiogram when analyzed by advanced computer methodology and the predictive accuracy of diagnostic criteria when implemented in ECG computer programs. Researchers are also investigating the use of well-documented populations and multivariate statistical techniques in designing new criteria. NHLBI and Framingham personnel and the European consortium (CSE) are collaborating on these studies.

LAS and LSM have developed statistical methods for determining the relative merits of diagnostic ECG statements. A paper was published in Computers in Cardiology using McNemar's test that allowed comparisons of ECG diagnostic statements reported in the literature to be tested for statistical significance; that paper made restrictive assumptions about Kendall correlation. In a follow-up paper published in the Journal of Electrocardiology the staff used data from the CSE study to refine estimates for the limits of Kendall correlation. Finally, a modification of McNemar's test involving the grading of ECG statements was applied to the CSE studies and presented at the 14th annual conference on "Computerized Interpretation of the Electrocardiogram."

LAS and LSM are also developing methods for comparing the relative diagnostic power of ECG algorithms using receiver operating characteristic (ROC) curves. Echocardiographic data from the Framingham Heart Study was used to determine the sensitivities/specifications for any value of the ECG-LVMI and the Cornell Voltage Index, thereby generating ROC curves for these two algorithms. A method for testing the difference of the ROC areas for significance was presented at the 14th annual conference on "Computerized Interpretation of the Electrocardiogram."

LAS is also involved in developing digital signal processing methods in electrocardiology. An important problem in ECG's is low frequency noise. LAS has developed null-phase digital filters with cut-offs at 0.5, 0.75, 1.0, and 1.2 Hz respectively. The performance of these filters in removing baseline wander without distorting the ST segment was compared to the standard 0.05 Hz filter and to the cubic spline method for removing baseline wander. The new methods were found to be superior to the commonly used technique. This study was reported at Computers in Cardiology. A refinement of the null-phase filters, a bilinearly transformed filter with a sharp cut-off at 0.8 Hz was described and its performance demonstrated at the 14th annual conference on "Computerized Interpretation of the Electrocardiogram."

**Future Studies**

The method of comparing areas under ROC curves using nonparametric statistics will be generalized using weights that assign degrees of membership to normal or abnormal classes. This will be applied to the echocardiographic and electrocardiographic data in the Framingham Heart Study.

The bilinearly transformed, null-phase method will be extended to a notch filter for power line interference and to a filter for high frequency noise.

### Interface for Computer Analysis of Molecular Databases

*Martha R. Horton, M.Sc.*
*with P.R. Krause, M.D. (NIAID/IRP/LCI); J.J. Bailey, M.D. (DCRT/LAS); E. Tyler, B.Sc. (DCRT/LAS)*

This project was developed to design and implement a much-needed interface between researchers using the various molecular biology databases and the computer programs that carry out analyses of these data. To use existing databases and programs, researchers must know the calling requirements of the various programs and the operating formats of the databases. This project will produce a computer software interface that will enable actual manipulation of the available data with a minimum of inconvenience to the researcher.

#### Research Progress in FY89

Preliminary design of the interface was performed by Dr. Philip Krause (NIAID), who developed the interface command language grammar. The commands of the grammar are designed to be readily understood by the researcher and to be easily amended or extended. This grammar has been implemented during FY89 by E. Tyler, under the supervision of Ms. Horton. The parser for the grammar is written in the C language for an IBM personal computer, but is readily adaptable to other computer systems. Additions or changes to the grammar can be easily made.

In order for a researcher to compare proteins or polynucleotides found in the laboratory with similar macromolecules from a molecular database, a computer program to search for the given sequences must be invoked by the interface program. In FY89, a comparison of search techniques was undertaken in order to determine the most appropriate technique to use with the interface.

In preparation for invoking various analysis programs, a survey was made of several existing DNA sequence analysis packages used on micro- and minicomputer systems.

LAS has acquired the HIV Sequence Database for Human Retroviruses and AIDS from the Los Alamos National Laboratory, and it is available for use in testing the interface.

#### Future Tools, Trends

The staff will implement the search algorithm(s) selected for the interface, and add the calling sequences for appropriate analysis programs to the interface. Debugging and enhancement of the various portions of the interface will be carried out and a software package with appropriate documentation will be produced. Progress will depend critically upon ability to retain staff that are currently in a temporary employment status.

### Computer Based Monitoring of Central Nervous System (CNS) Function in the Neurosurgical Intensive Care Unit

*Martha R. Horton, M.Sc.*
*with R.C. Burgess, M.D., Ph.D. (Cleveland Clinic Foundation, Department of Neurology)*

Evaluation of the integrity and function of the CNS in critically ill patients with ischemic, metabolic, or traumatic brain insults is ordinarily assessed by episodic neurologic exams and cannot be done in an ongoing fashion. The purpose of this project is to develop and test a computer-based system for acquisition, analysis, and display of scalp-recorded neuroelectric signals (electroencephalogram and evoked potentials). This tool can be used to investigate the degree of dysfunction in neurologically

impaired patients, to correlate the indices developed with other measures of cerebral function, and to evaluate the effectiveness of various therapeutic interventions. The system automatically sequences from one stimulus modality to the next while simultaneously adjusting preamplifier/filter parameters, generating cumulative displays, and saving the results for trend analysis. LAS has implemented programs for digital filtering, frequency domain analysis, and displays using color to highlight important trends.

**FY89 Progress**

.The central nervous system (CNS) monitoring system developed by LAS from FY85–87 was used in four projects during FY89.

1) A pilot study was conducted on continuous multi-modality CNS monitoring of Neurosurgical Intensive Care Unit (NICU) patients at The Cleveland Clinic Foundation, Cleveland, Ohio.

The original system was designed to allow automatic sequencing through a protocol of evoked potentials and EEG's, and adjustment of amplifier settings and display, analysis, and stimulation parameters. The current pilot study is the first application of the system involving human patients.

The CNS monitoring system was modified in preparation for use in the NICU. Special attention was paid to refining data collection and presentation in order to streamline operation and to optimize utilities for monitoring NICU patients. The modifications include addition of a Nicolet SM2000 amplifier, and creation of a screen-oriented interface to simplify protocol formation, editing, and data review. These functions cannot be carried out by noncomputer experts.

The pilot study is designed to identify those comatose patients most likely to benefit from continuous CNS monitoring, and also to identify the most useful methods of monitoring protocols, data analysis, and data presentation.

Initial phases of the pilot study have uncovered some clinical logistical problems inherent in NICU monitoring that must be addressed in a system solution. Merely placing the electrodes on the patient's head may cause some elevation of the ICP; ventriculotomies, ICP monitors, and surgical openings compete for scalp space; and tests such as MRI scans require removal of the electrodes. Solutions to these problems will be minimized by use of on-call monitoring personnel and the establishment of routine procedures.

2) Analysis was conducted of EEG using dynamic state space analysis.

Software developed on the CNS system in FY89 was used to perform nonlinear analysis of EEG data, including the calculation of Hausdorff dimensions. Presentation graphics were developed to allow N-dimensional display and interactive thresholding, and to facilitate clustering using chaos theory techniques.

3) Adaptation of time varying filter techniques were developed to time locked signal averaging of clinical evoked potentials.

Existing software, previously used for optimal filtering of simulated time locked signals, was prepared for use on evoked potentials. A program to convert data into the correct form for time varying filter manipulations has been completed.

4) Software was developed that performs pharmacokinetics analysis for optimal dosing of anticonvulsant medications.

The CNS system was used in development of a program that facilitates the calculation of

pharmacokinetic variables for studies of drugs with linear elimination characteristics. Following the administration of a single dose of anticonvulsant, serum levels are measured with time, then fit to a mathematical model in order to predict the optimal maintenance dose for an individual patient.

**Future Tools, Trends**

The staff used results of the current pilot study to guide development of additional signal processing and display utilities, such as providing a trend graph of evoked potential peak latencies, EEG power asymmetry, etc. A study will be run in the NICU to evaluate the clinical effectiveness of multimodality CNS monitoring.

**Analysis of Physiological Signals**

*Erik W. Pottala, Ph.D.*
*with J.J. Bailey (DCRT/LAS), J.A. Dvorak (NIAID/LPD), and K. Rasmussen (NICHD/LCE)*

This project involves the development and application of minicomputer and PC-based signal processing techniques for analysis of physiological signals, e.g., electrocardiogram, electromyogram, and electroencephalogram. The LAS minicomputer-based system provides a general purpose analog to digital conversion facility and an ability to pre- or postfilter the signals with a variety of analog and digital techniques.

Investigators at the Laboratory of Comparative Ethology, NICHD, are studying heart rate variability (HRV) in free-living simians. In FY89 methods were developed to deal with muscle noise, artifact, extrasystoles, and other confounding factors that may interfere with obtaining an accurate measure of HRV in the frequency domain. Currently, the analog to digital conversion and all analysis of the data are being implemented on an IBM/XT using Microsoft QuickBasic-4.

The development of new filtering techniques is continuing on the IBM/XT using MATLAB, a matrix manipulating software package. In FY89 new filters for preprocessing ECG data to suppress baseline wander without distortion of the ST segment were designed and tested against other commonly used methods. Reports on these studies were presented at Computers in Cardiology and at the 14th annual conference on "Computerized Interpretation of the Electrocardiogram." MATLAB on the IBM/XT permits remarkable facility for interactive design of digital processing techniques and display of results.

In FY89 acquisition of an Apple Macintosh II microcomputer system with a Spectral Innovations signal processing board was begun. This system will replace the obsolete LSI-11 based minicomputer system.

**Future Tools, Trends**

In FY90 nearly all signal processing activities will be transferred to a new Macintosh II system. This is a state-of-the-art system that can handle large data arrays and has excellent display and interactive capabilities. With the signal processing board, it has enhanced computational speed (e.g., can perform real 1024 FFT in 3.4 msec.). This system will be able not only to perform high speed analog to digital conversion, but also to do additional signal processing that is presently done on the DECsystem-10.

**Computer Systems for Nuclear Medicine**

*Margaret Douglas, B.A.*
*with: J.J. Bailey, M.D. (DCRT/LAS); S.L. Bacharach, Ph.D., M.V. Green, M.S. (CC/Nuclear Medicine); R.O. Bonow, M.D. (NHLBI/CB)*

LAS develops systems for computer-based mathematical analysis, pattern recognition, and image processing in support of diagnostic activities in the Clinical Center Nuclear Medicine Department and collaborating institutes. Diverse applications include the following: parameters of ventricular function extracted from radionuclide ventriculography; tumor detection and imaging using radio-labeled monoclonal antibodies and functional imaging of the central nervous system using PET scan technology; and the correlation of function as obtained from PET scans with structure as revealed by CT and MRI scans.

**FY89 Progress**

LAS in collaboration with the Clinical Center Nuclear Medicine Department has, over the past three years, designed and specified a general purpose image processing system, MIRAGE. Programming was performed by contractors monitored by LAS and the Nuclear Medicine Department. The completed basic system has been ported to other NIH computer systems. This year three-dimensional imaging has been added to MIRAGE.

LAS in collaboration with the Nuclear Medicine Department is evaluating various general purpose image processing facilities for the Nuclear Medicine Department, including a Sun 4 Workstation with TAAC-1 hardware and software, and the Mayo Clinic image processing package ANALYZE. LAS and the Nuclear Medicine Department are designing modifications and additions for the TAAC-1 system to make it more responsive to the needs of the Nuclear Medicine Department. Both the TAAC-1 and ANALYZE systems are being used to investigate problems of advanced three-dimensional visualization of medical images.

Other applied image processing activities include LAS participation in a committee that is developing guidelines for general purpose image processing software and user interface standards for use on a wide variety of computer systems at NIH; LAS collaboration in an effort to determine parameters of cardiac volume and radius of curvature from echo data; investigation of the use of neural networks in the problem of image segmentation and classification; campus-wide consulting on image processing, especially in the area of PC-based image processing; serving as image processing representative to groups involved with NIH campus computer connectivity; and investigation of PC CAD. LAS personnel also sponsor seminars through the NIH-Wide Image Processing Group and collaborate in publishing the Directory of NIH Image Processing Facilities.

**Future Tools, Trends**

The Nuclear Medicine Department has acquired several systems that collect three-dimensional data. The most recent is the Posicam scanner capable of acquiring 21 simultaneous slices per second, each slice consisting of 256 x 256 pixels. Therefore, LAS will continue to investigate three-dimensional visualization, multimodality registration, and development of fast interactive algorithms for analysis of large volumes of data. The Clinical Center has set up a committee and asked LAS to take an active role in the design of an image transmission network; design and specification of at least two levels of standards for image processing workstations; and development of a consensus approach to medical image archiving issues.

**Computer-based Analysis and Image Processing in Electron/Light Microscopy and X-ray and Electron Energy Spectroscopy**

*Margaret Douglas, B.A.*
*with D. Johnson, M.D. (Children's Hospital)*

This project is directed toward the development of computer-based mathematical and statistical analyses, pattern recognition, and image processing of data, principally x-ray micrography and electron energy loss spectra, and the electron/light microscopy images of biological specimens.

A collaborative project with the Department of Neurosurgery at Children's Hospital has been completed. This project involved analysis of micrographs of rat spinal cord to determine parameters of injury (i.e., loss of axon volume, number of axons, inflammatory infiltrate, etc.). The system developed at LAS was duplicated so that both LAS and Children's Hospital have compatible systems for data analysis.

Other applied image processing activities include investigating automated cell classification from micrographs and review of the field of medical image processing of micrographs for an invited journal review article.

Future efforts will include the development of a new Macintosh-based image and signal processing system.

## Publications and Presentations


Bailey, J.J.; Campbell, G.; Horton, M.R.; Shrager, R.I.; Willems, J.L.: Determination of Statistically Significant Differences in the Performance of ECG Diagnostic Algorithms: An Improved Method. *J. Electrocardiology* 21 Suppl:S188-S194, 1988.

Bailey, J.J.; Campbell, G.; Horton, M.R.; Shrager, R.I.; Willems, J.L.: How Can Statistically Significant Differences in the Performance of ECG Diagnostic Algorithms Be Determined? *Computers in Cardiology.* No. 87CH2544-5. IEEE Computer Society, Silver Spring, Maryland, 1988, pp. 35–39.

Boverman, J.F.; Burgess, R.C.; Bourgeois, B.F.D.; Turnbull, J.P.; Horton, M.R.: On-line Anticonvulsant Pharmacokinetics Analysis on a Multi-User System. Submitted to the 11th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, September, 1989.

Burgess, R.C.: ICU Monitoring. American EEG Society Annual Course Syllabus, San Diego, CA, October, 1988.

Burgess, R.C.: Neurophysiologic Monitoring in the NICU. Neurosurgical and Neurological Intensive Care Course, The Cleveland Clinic Foundation, May 1989.

Campbell, G.; Douglas, M.A.; Bailey, J.J.: Nonparametric Comparison of Two Tests of Cardiac Function on the Same Patient Population Using Entire ROC Curve. *Proceedings of the Computers in Cardiology Conference* (in press).

Campbell, G.; Levy, D.; Lausier, A.; Horton, M.R.; Bailey, J.J.: Nonparametric Comparison of Entire ROC Curves for Computerized ECG Left Ventricular Hypertrophy Algorithms Using Data from the Framingham Heart Study. *J. Electrocardiology* 1989 (in press).

Campbell, G. and Shrager, R.I.: Robustness of Weighted Estimators of Location: A Small Sample Study. *Computing Science in Statistics: Proceedings of the 20th Symposium on the Interface,* E.J. Wegman, (Ed.), pp. 650–55, 1988.

Douglas, M.A. and Trus, B.L.: An Introduction to Image Processing in Medical Microscopy. *Medical Progress through Technology* (in press).

Fletcher, J.E.; Shrager, R.I.; Bailey, J.J.; Jackson, W.L.: A Kinetic Model of T Lymphocyte Interactions with HIV. Presented at the Workshop on Modeling the Interaction of HIV with the Immune System, Santa Fe Institute, October 5, 1988.

Fujimori, K.; Covell, D.G.; Fletcher, J.E.; Weinstein, J.N.: A Modeling Analysis of the Global and Microscopic Distribution of IgG, F(ab')2, and FAB in Tumors. *Cancer Research* 49, pp. 5656–5663, October 15, 1989.

Fujimori, K.; Weinstein, J.N.; Covell, D.G.; Fletcher, J.E.: An Analysis of IgG, f(ab')2, and Fab Percolation Through Tumors. *Jour. Soc. of Nuclear Med.* June, 1988.

Gradwohl, J.R.; Pottala, E.W.; Horton, M.R.; Bailey, J.J.: Comparison of Two Methods for Removing Baseline Wander in the Electrocardiogram. Computers in Cardiology. IEEE Computer Society, Silver Spring, Maryland 1989 (in press).

Lewis, N.; Levin, I.W.; Shrager, R.I.: Applications of Fourier Transform Raman Spectroscopy to Highly Fluorescent Membrane Assemblies: Polyene Antibiotic/dipalmitoylphosphatidylcholine Interactions. Presented at the Biophysical Meeting, Cincinatti, Ohio, Feb. 13–16, 1989.

Pottala, E.W.; Bailey, J.J.; Horton, M.R.; Gradwohl, J.R.: Suppression of Baseline Wander in the ECG Using a Bilinearly Transformed, Null Phase Filter. *J. Electrocardiology,* 1989 (in press).

Richards, E.W.; Hamm, M.W.; Fletcher, J.E.; Otto, D.A.: The Binding of Palmitoyl-CoA to Bovine Serum Albumin. *Journal of Biological Chemistry,* 1989 (in press).

Tyler, E.: A Look at Several DNA Sequence Analysis Packages for Use on Micro- and Mini-Computers. Independent Research Study, Duke University, Durham, North Carolina.

Willems, J.L.; Campbell, G.; Bailey, J.J.: Progress on the CSE Study—Applications of McNemar's Test Revisited. *J. Electrocardiology* 1989 (in press).

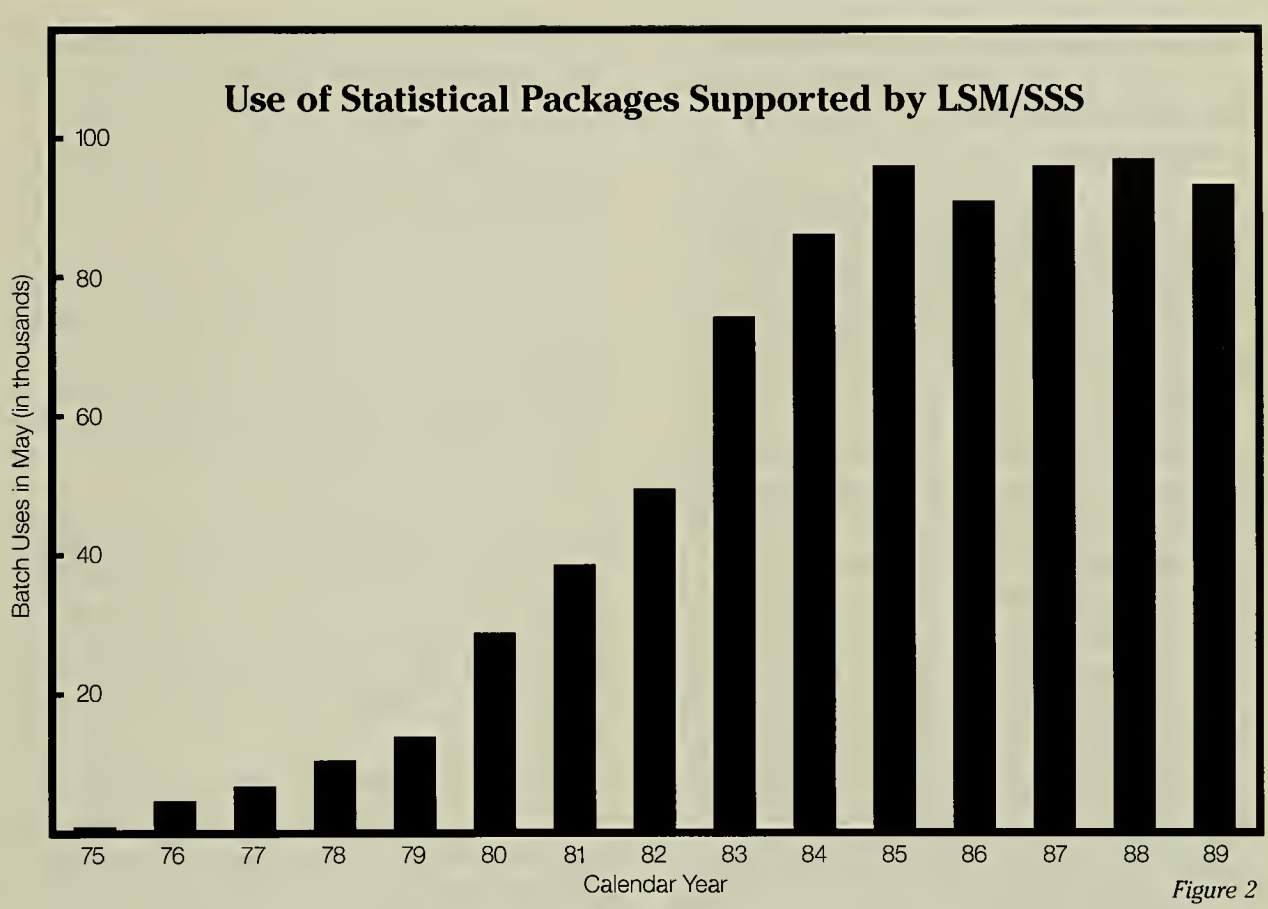



*Figure 2*

sessions, 600 GRAPH sessions, and 250 VMAP accesses per month, down to relatively few sessions for specialized programs such as REDUCE, DNALAB, GLIM, and CART. A new version of DNAdraw for IBM compatible microcomputers, which formats, annotates, and displays DNA sequences for publication, was nearing completion as FY89 ended.

## Consultation and Collaboration in FY89

LSM provides consultation in a wide range of scientific fields, and interpretation of the results of statistical and other scientific computation. Staff members also provide consulting in statistics, biomathematics, and computer analysis of DNA sequences. Consultations can be brief or extensive depending on the complexity of the statistical or scientific research. Some evolve into long term collaborations and scientific publications. These collaborative efforts include the following:

- Dr. G. Cutler (NICHD/DEB). Multiple consultations were held with members of Dr. Cutler's Developmental Endocrinology Branch. These dealt with treatment, detection, and prevention

of childhood and juvenile endocrinology problems. Almost all of the studies involved repeated measures analyses suitable for the study of longitudinal data.

- Dr. R. Callahan (NCI/LTIB). Statistical methods for investigating the association of the amplification of the oncogene c-erbB2 in human breast tumors with lymph node number and survival are being studied. A joint paper is in preparation.

- Dr. A. Cheever (NIAID/LPD). Schistosomiasis infections are studied in different strains of mice by different parasite strains. Egg destruction in tissues under different treatment regimes was studied. LSM staff performed the statistical analysis.

- A. Mitz (NIMH/LNP). Single-unit activity was studied in the premotor cortex of rhesus monkeys while they learned new visuomotor associations. Models based on Poisson processes were used to study the learning behavior of the monkeys using neuronal activities. LSM assisted with statistical analysis.

- Dr. D. Fraser (NIAMS/ARB). Patients are rated zero to three with regard to muscle inflammation from polymyositis by three different scales—clinical impression (the gold standard), MRI, and muscle biopsy. The latter two scales are compared with the clinical impression with regard to sensitivity and specificity. Statistical techniques that were applied include McNemar's test and Wilcoxon's signed rank test for paired data. In addition, ROC curve analyses were studied as an application. LSM assisted with statistical analysis.

- Dr. G. Straw, Dr. M. Jaffe (NIMH/MHIRP). A study of the electro-retinogram amplitude response of the dark-adapted alpha-wave of human subjects to different light intensities has been analyzed. LSM provided statistical expertise concerning the modeling of the data with Naka-Rushton and Hill equations using weighted least-square techniques.

- Dr. J.N. Weinstein (NCI/DCBD/LMB). The study of dipyridamole (DPM) as an inhibitory of azidothymidine (AZT) and 2',3'-dideoxycytidine (ddC) against human immunodeficiency virus (HIV-1) in human monocyte/macrophages was pursued. Cells from healthy humans were infected with an HIV-1 isolate; these cells were treated with combinations of DPM and AZT and the viral growth was measured. LSM provided nonparametric analyses based on Jonckheere's test to assess any monotonic trends in viral growth for increasing concentrations of AZT and/or DPM.

- J. Angle, S. Tokarick (Econometric Research Service, USDA). A problem of reliability analysis required solution of a linear system of equations with symbolic formula coefficients. LSM assisted in formulation of the problem for analysis by the MACSYMA program on the ALW local area network.

- Dr. E. Ungar (NHLBI/CB). Aspects of the human circulatory system are modeled by network equations similar to Kirkhoff's laws for electrical circuits. LSM assisted in solution of a derived system of symbolic equations, using the MACSYMA program on the ALW local area network.

- Consultations on the use of the DNAdraw program for producing publication-quality drawings of DNA sequences included Dr. S. Cheng, (NCI/DCBD); Dr. R. Mage, (NIAID/LI);

Dr. P. Qasba, (NCI/DCBD); Dr. S. Tronick, (NCI/DCE/LCMB); Dr. R. Wickner, (NIDDK/LBP).

- Consultations on the use of the GRAPH program for producing publication quality graphs were numerous and included consultations with J. Guroff (NIMH/BP); Dr. A. Minton (NIDDK/LBP); Dr. W. Stewart (NIDDK/LAC); Dr. J. Zimmerberg (NIDDK/LMB).

Consulting services will continue with available staff. Due to staff reductions, biomathematics consultation will be very limited.

## Research and Development in FY89

Research projects in LSM vary widely; they include statistical methodologies for biomedical applications, language processing for medical information systems, scientific computer printing and graphics, and research in mathematical and statistical methods.

During FY89, SMS staff were active in mathematical statistical research in a variety of areas. Studies of size and shape variables continued to provide methods for studying random proportions or ratios of common occurrence in biomedical data. Studies were directed largely to properties of patterned covariance matrices that occur commonly in biological data. Several data analyses using these results are in progress. One paper on normalizing ratios and proportions was completed.

Investigations in nonparametric statistics continued. A study of the robustness of weighted estimators of location using computer simulations has appeared in print. Studies of methods for the analysis of correlated proportions were continued and several presentations based on applications to ECG program evaluation were made, all in the United States.

A comprehensive study of Receiver Operating Characteristic (ROC) curves that began in FY88 was continued. The ROC curve is an important tool in the evaluation of diagnostic biomedical tests. Development of a nonparametric approach was continued, with the main problem remaining the comparison of two diagnostic ROC curves in the important situation in which the same patients are used for both tests. Presentations on this topic were made at national and international meetings (all in the U.S.A.).

The study of the problem of optimal unbiased estimation of variance components in mixed models with possibly unbalanced data was continued, with some implementation of computer programs for testing designs. Similar mathematical methods were used to study statistical inferences in physical situations where quantum mechanical methods are needed; particularly with respect to statistical problems of estimation that arise in modern biotechnology (PET and bioluminescence assay techniques). A number of invited lectures on these topics were given in the past year.

Biomathematics and Computer Science section research on the algebra of rings and modules continued. During the year, a method for expressing relation categories of modules as universal algebras was developed.

During FY89, the staff of the Biomathematics and Computer Science Section spent considerable time upgrading the local area network facilities available to LSM staff. The current configuration includes six PS/2 Model 80's, four PS/2 Model 60's, two PC/AT's and five 3Com3 Station microcomputers connected to the PCB

local area network, one Macintosh II and one Sun 3/50 workstation to be connected to the advanced laboratory workstation (ALW) local area network. The current LSM server was upgraded, and procurement of a second LSM server is in progress. In addition, a substantial amount of commercial software such as MACSYMA (on the ALW), Mathematica and Publisher's Type Foundry has been adapted to serve LSM needs in scientific research, consulting, and software development. Laser printing capability was increased, and printing utilities designed for NIH needs were upgraded.

In the Medical Information Science Section, research in medical linguistics continued at the same level. A parsing methodology based on a predicate calculus formalism is being implemented. A separate project for development of a comprehensive lexicographic data base will be continued, and used as a system facilitating collaborative research. LSM continued collaboration with the Laboratory of Pathology, NCI/DCBD/LP, and DCRT/DMB, using LSM's automatic encoding system to maintain and improve the surgical pathology report data base of the Clinical Center. In a number of instances, MIS staff assisted in the formulation of special queries for this data base.

## Future Plans: Software Development, Contracts, Network Development

Laboratory service and research will be maintained as availability of staff permits. LSM's high level of support for IBM S/370 statistical software systems will be continued. Contract support for MLAB will continue at a reduced level. Support of GRAPH, DNAdraw, and VMAP for use on the PCB network will continue. Further development of scientific printing software is anticipated.

LSM will continue its contribution to software testing and development for two DCRT local area computer networks: A network supporting IBM compatible and Macintosh microcomputers, and the advanced laboratory workstation project of CSL supporting UNIX-based workstations. Support of SAS in a PC/networking environment will also continue.

## Research Projects

### Automated Data Processing of Medical Language

*G. Dunham*
*with A.W. Pratt (DCRT/OD), M.G. Pacak (Guest Researcher), S. Harper (DCRT/LSM), E. Jaffe (NCI/DCBD)*

The major objective of the project is the development of methods for the automatic processing of natural medical language. Research in medical linguistics included the development of paraphrasing rules for medical noun phrases, a semantic grammar, and a generalized morphological analysis approach suitable for general language and data compression, as well as medical language morphology. Development of the Lexicographic Environment Software, a platform for medical lexicography and the development of medical language processing continued.

Research focused on the theoretical and developmental issues of integrating medical text and the existing theorem prover for relevance/entailment predicate calculus with the Lexicographic Environment Software. For example, text sentences may have a logical equivalent or explication, which is in turn integrated with the

lexical aspect of the system. Also, the optimal way to store the continuum of text, from morpheme to discourse, was investigated. All components are designed to facilitate the adjoining of learning procedures to bear some of the task of developing detailed medical dictionaries and language processing procedures.

Collaboration continued on the Clinical Information Utility with the Laboratory of Pathology, NCI/DCBD/LP, and with DCRT/DMB, in using LSM's automatic encoding system to maintain and improve the data base of Clinical Center surgical pathology reports. Progress was made on obtaining and automatically encoding NIH surgical pathology data both from the Clinical Center Medical Information Science files, and from files covering the period 1959–1977.

**Computer Graphics and Applications**
*M.B. Shapiro*

The main objective of this project is the application of computer graphics and related methods to NIH research problems.

Work continued on the IBM/PC versions of the GRAPH and DNAdraw programs, which are still heavily used on the DECsystem-10. The PC/GRAPH program is now about 50 percent complete, while the PC/DNAdraw program is 80 percent complete and ready for preliminary use and testing. PC/DNAdraw is a menu-driven program that outputs annotated and highlighted DNA sequences directly to the HP Laser Jet printer. It contains features that take advantage of the Laser Jet such as built-in fill and shade patterns.

**Discrete Mathematics and Applications**
*G. Hutchinson, Ph.D.*

The project objective is to develop mathematical theory and computational techniques using discrete mathematics (algebra, combinatorics, and graph theory), and to apply such methods to appropriate problems of biomedical research and computer science.

Work continued on the characterization of rings that possess the same associated lattice theory of modules. A method was devised for expressing associated relation categories within universal algebra. A publication appeared during the year, and a book is in progress.

Development of the Scientific Printing Utility (formerly called VMAP) was delayed by the failure to obtain any bids on an advertised procurement for public domain computer fonts in FY88. A contract procurement partially meeting SPU requirements for public domain computer fonts was awarded during FY89, and is now in progress. SPU public domain font requirements will also require BCS section efforts, which are currently in progress. SPU is currently being tested in configurations using both commercial and public domain computer fonts.

Analysis of lattices and categories of modules will continue. A new initiative in the use of algebraic methodology in computer software technology is in preparation. Development of scientific printing software methodology will continue.

### Multivariate Statistical Analysis

*J.E. Mosimann, Ph.D.*
*with G. Campbell (DCRT/LSM), M.V. Ratnaparkhi (Wright State University, Dayton, Ohio), Paul Pirlot (Universite de Montreal), E. Depiereux (Facultes Notre Dame de la Paix, Namur, Belgium).*

The objective of this project is the study of multivariate ratios or proportions. Studies of random proportions and ratios, including those that follow a multivariate lognormal distribution as well as those following the distribution of a mixture of Dirichlet distributions, were continued. In the past year, the major emphasis was on obtaining closed form solutions for the eigenvalues and eigenvectors of certain patterned matrices that arise frequently with biological data. The patterns are those of "persymmetric" (includes Toeplitz) matrices, as well as a pattern that reflects a single "unique or dominant" variable related to any number of variables equicorrelated among themselves. In addition, a paper with M.V. Ratnaparkhi on normalizing transformations of ratios and proportions was completed. These research results are being assembled for a research monograph on the subject.

### Algebraic Methods in Statistics

*J.D. Malley, Ph.D.*

This continuing project will develop new methods in statistics, using algebraic techniques, that have applications for biomedical research.

A continuing effort toward a solution to the problem of statistical estimation in the presence of missing data, or data having a known patterned covariance matrix, has recently culminated in the solution of both these problems for multinormally distributed data, as well as for counted data (multiway contingency tables).

The method solves for the maximum likelihood estimate of the parameter vector, thereby generating estimates that are unbiased as well as asymptotically fully efficient. The technique invokes the Dempster-Laird-Rubin (1977) EM algorithm and Jordan algebras to iteratively solve for the parameter vector. Known properties of Jordan algebras are invoked to guarantee convergence, so that each iteration only uses a matrix multiplication by a single fixed matrix with an order that is, at most, twice the number of parameters involved.

Hence, convergence is guaranteed in a nearly closed form solution to, for example, repeated measure designs with missing data, to variance component estimation (with estimates that are guaranteed to be positive), to variance-covariance problems such as those that occur in genetics and time-series problems, and to patterned covariance estimation such as those that occur in structured testing and longitudinal and growth model problems. Again, missing data is allowed, without the need to drop any incomplete data records, in any of the problems just described.

Some of the results are already contained in a paper under review, and the complete solution will be incorporated into the final revision.

### Quantum Statistical Inference

*J.D. Malley, Ph.D.*
*with J. Hornstein, Ph.D. (Naval Research Laboratory)*

This interdisciplinary project is concerned with the theory and biomedical applications of recently developed statistical decision procedures. This methodology allows the statistician, for the first time, the opportunity to undertake nearly classical statistical analyses that are fully

consistent with the results and first principles of quantum mechanics and quantum chemistry.

The procedures involve the outlook of classical statistical decision theory (estimation, classification, inference, etc.) and are worked out within the ab initio calculation framework of quantum mechanics. A first, specific application of this method is the quantum signal detection (photon counting) problem of realtime, bioluminescent imaging, a technique that uses semiclassical imaging with user-defined (ad hoc) thresholds. The fully quantum mechanical solution is, however, known in principle to be optimal, making the most physically and statistically efficient use of the given biomedical data.

Other applications could include fully efficient PET scan data imaging and analysis, as well as problems in the quantum chemistry of biologically active molecules, their structure, and reactivity.

### Nonparametric Statistics

*G. Campbell, Ph.D.*

A comprehensive study of the theory and application to biomedical research of Receiver Operating Characteristic (ROC) curves has continued. A theory of nonparametric estimation of the area under the ROC has been developed. This can be used to compare two diagnostic tests via ROC area when the data are ordinal categories rather than continuous variables. Also, the theory of ROC curves has been extended to fuzzy data; instead of knowing that a patient is diseased or nondiseased, it is sometimes useful to view the patient's disease state to be a continuous variable between zero and one, where zero indicates normal and one indicates diseased. With such data, one can define weighted estimates of sensitivity and specificity and hence the fuzzy ROC curve. The area under the ROC can be calculated easily. This avoids the issue of bias, which is usually introduced when the two groups are not well separated and individuals without clear identification are dropped from the analysis. Work is being extended to the statistical properties of analyses for comparing areas under two fuzzy ROC's, as well the important problem of restriction to specific sections of the ROC curve (for example, limiting to specificities greater than .5). A study of the Lomax and Pareto distributions with application to the problem of parameteric ROC modeling of data from the same patients of diagnostic tests has begun.

## Publications and Presentations


Ali, I.U.; Campbell, G.; Merlo, G.R.; Smith, G.H.; Callahan, R.; Lidereau, R.: Multiple genetic alterations in human breast cancer and their possible prognostic significance, *Cancer Cells,* Vol. 7: Molecular Diagnostics of Human Cancer, Cold Spring Harbor, NY, 1988 (in press).

Bailey, J.J.; Campbell, G.; Horton, M.R.; Shrager, R.I.; Willems, J.L.: Determination of statistically significant differences in the performance of ECG diagnostic algorithms: an improved method. In: Kligfield, P.; Bailey, J.J. (Eds.), Computerized Interpretation of the Electrocardiogram XIII, *J. of Electrocardiology,:* 188–192, 1988.

Bailey, J.J.; Campbell, G.; Horton, M.R.; Shrager, R.I.; Willems, J.L.: How can statistically significant differences in the performance of ECG diagnostic algorithms be determined? *Computers in Cardiology,* 35–39, 1988.

Campbell, G.; DeLeo, J.M.: Fundamentals of fuzzy receiver operating characteristic (ROC) functions. In: Berk, K. (Ed.), Computing Science and Statistics: *Proceedings of the Twenty-First Symposium on the Interface,* 1989 (in press).

Campbell, G.; Douglas, M.A.; Bailey, J.J.: Nonparametric comparison of two tests of cardiac function on the same patients using the entire ROC curve. *Computers in Cardiology,* 1988 (in press).

Campbell, G.; Levy, D.; Lausier, A.; Horton, M.; Garrison, R.J.; Bailey, J.J.: Nonparametric comparison of entire ROC curves for computerized ECG left ventricular hypertrophy algorithms using data from the Framingham Heart Study. In: Kligfield, P., Bailey, J.J., (Eds.), Computerized Interpretation of the Electrocardiogram XIV: *J. Electrocardiology* 22, 1989 (in press).

Campbell, G.; Shrager, R.I.: Robustness of weighted estimators of location: a small-sample study. In: Wegman, E.J. (Ed.), *Computing Science and Statistics: Proceedings of the Twentieth Symposium on the Interface* 650–655, 1988.

Campbell, G.: Common Misuses of Chi-Square Statistics. Laboratory of Tumor Immunology, National Cancer Institute, Bethesda, Maryland, April, 1988.

Campbell, G.: Evaluation of Diagnostic Tests: Statistical Techniques with Applications to ECG Analysis. Division of Medical Informatics, University of Leuven, Leuven, Belgium, July, 1988.

Campbell, G.: The Use of ROC Curves in the Comparison of Diagnostic Tests. Centre Informatique, University of Liege, Liege, Belgium, July, 1988.

Campbell, G.: Nonparametric Statistical Methods for the Comparison of Two Diagnostic Tests Using ROC Curves from Ratings Data. XIVth International Biometric Conference, Namur, Belgium, July, 1988.

Campbell, G.: Nonparametric Comparison of Two Tests of Cardiac Function on the Same Patients Using the Entire ROC Curve. Computers in Cardiology, Bethesda, Maryland, October, 1988.

Campbell, G.: Fundamentals of Fuzzy Receiver Operating Characteristic (ROC) Functions. Twenty-first Symposium on the Interface of Computer Science and Statistics, Orlando, Florida, April, 1989.

Campbell, G.: Nonparametric Statistical Methods for the Comparison of Two Diagnostic Tests Using ROC Curves from Ratings Data. XIVth International Biometric Conference, Namur, Belgium, July, 1988.

Campbell, G.: Nonparametric Comparison of Two Tests of Cardiac Function on the Same Patients Using the Entire ROC Curve. Computers in Cardiology, Bethesda, Maryland, October, 1988.

Campbell, G.: Fundamentals of Fuzzy Receiver Operating Characteristic (ROC) Functions. Twenty-First Symposium on the Interface of Computer Science and Statistics, Orlando, Florida, April, 1989.

Fuller, K.R.; Hutchinson, G.: Exact embedding functors and left coherent rings, *Proc. Amer. Math. Soc.* 104: 385–391, 1988.

Malley, J.: (1) Some uses of higher algebra in quantum chemistry with applications to molecular biology, (2) quantum statistical inference, (3) some recurrent problems in biomedical data analysis and a discussion of some of their solutions. Jointly sponsored by the Depts. of Chemistry, Mathematics and Statistics, North Dakota State University, October 19–22, 1988.

Malley, J.: Jordan algebras and quantum statistical inference and estimation, with possible applications, Dept. of Mathematics, University of Virginia, December 2, 1988.

Malley, J.: Algebraic methods in statistics, Dept. of Mathematics, University of Maryland, College Park, February 4, 1989.

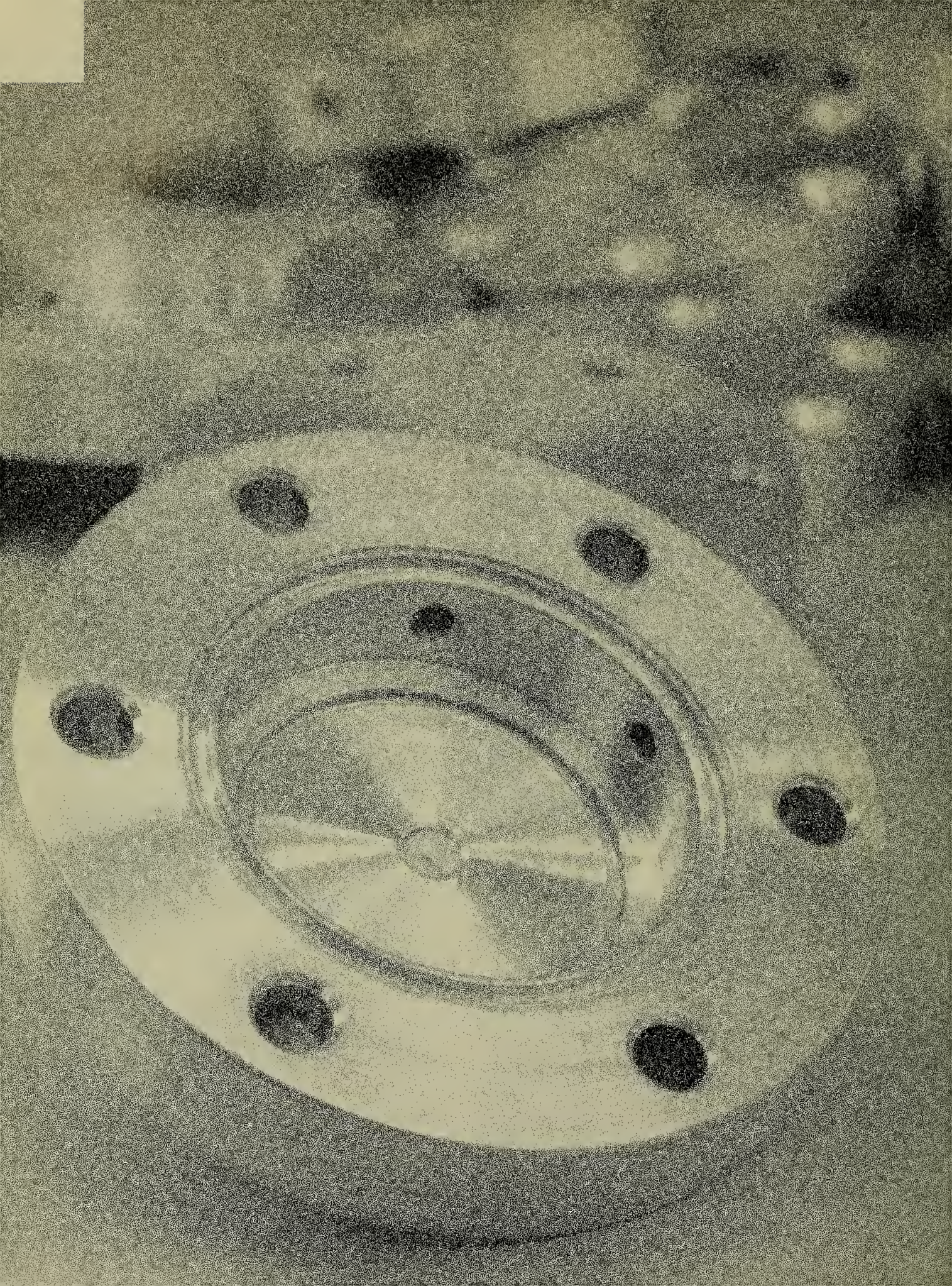
Malley, J.: Algebraic methods in statistics, Dept. of Mathematics, University of Maryland, Baltimore Campus, February 8, 1989.

Malley, J.: Quantum statistical inference with applications to biotechnology, Naval Research Laboratory, Washington, DC, March, 1989.

Malley, J.: Variance and variance-covariance component estimation, Dept. of Statistics, Oregon State University, April 18–19, 1989.

Senapathy, P., Shapiro, M.B., Harris, N.L.: Splice junctions, branch point sites and exons: sequence statistics identification, and applications to the genome project. In: Doolittle, R. (Ed.), Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences. *Methods in Enzymology,* Academic Press, 1989.

Willems, J.L.; Campbell, G.; Bailey, J.J.: Progress of the CSE diagnostic study—application of McNemar's test revisited. In: Computerized Interpretation of the Electrocardiogram XIV: *J. of Electrocardiology* 22 (in press).



# Physical Sciences Laboratory

*George H. Weiss, Ph.D., Chief*

The activities of the Physical Sciences Laboratory encompass a wide range of scientific disciplines, which include applied mathematics, biophysics, molecular graphics, physical chemistry, and physics. The laboratory has eight staff members, including post-doctoral fellows, but during the course of the year there are a large number of visitors who perform collaborative research with the permanent staff. A number of the projects include both theoretical and experimental components.

Examples of some current research projects in the Physical Sciences Laboratory include the use of Nuclear Magnetic Resonance (NMR) techniques to detect the influence of membrane lipids on water. These measurements, when completed, will be related by directly measured forces between membranes. Measurements of the change in molecular disorder that accompanies changes in molecular interactions as a function of the separation of membranes have allowed PSL researchers to develop a model of molecular motion in the field of force as a distribution of neighboring molecules. In other work, the theory of photon migration in turbid media, developed by members of the laboratory, has been successfully extended to the interpretation of optical absorption spectroscopy experiments carried out at the University of Pennsylvania. Many aspects of the theory remain to be developed for application to measurements made on a variety of human tissues. Further results have been obtained in the theory of the diffusion of geminate pair molecules in the presence of different types of possible binding mechanisms. The theory is currently being applied to the interpretation of experiments on proton dissociation reactions.

PSL staff members have been invited to give seminars on their work at many universities and have been invited to give lectures at a number of national and international symposia. Dr. B.K. Lee presented many seminars and lectures in Korea during the past year on his work in molecular graphics and the theory of hydrophobicity. Dr. R.J. Nossal was an invited speaker at a symposium on optical properties of mammalian tissues sponsored by the Optical Society. Dr. V.A. Parsegian gave a number of invited seminars in Russia, as well as being an invited speaker at a symposium in memory of Alex Mauro at the Rockefeller University, at a Gordon Conference on the Chemistry of Supramolecules and Assemblies, and at a symposium held at MIT. Dr. G.H. Weiss gave a number of invited seminars in Israel, and was an invited lecturer at a Gordon Conference on Statistics in Chemistry and Chemical Engineering.

## Research Projects

### Diffusion Influenced Reactions

*Noam Agmon, Ph.D.*
*with D. Huppert, Ph.D. (Tel Aviv University); M. Maroncelli, Ph.D. (Pennsylvania State University); A. Szabo, Ph.D. (NIDDK/LCP); G.H. Weiss, Ph.D. (DCRT/PSL); R. Zwanzig, Ph.D. (NIDDK/LCP).*

This project involves the investigation of different problems that arise in the theory of reactions when the rate limiting step is due to diffusion. This theory is important for interpreting experiments on the kinetics of chemical reactions in solution. The standard theoretical development requires a reaction to take place when two diffusing molecules come into contact. In reality, a reaction can be much more complicated. These complications affect the kinetics of

the reaction. Work on this project consisted of an elucidation of a number of these complicating features.

The first of these was an extension of the theory of reversible diffusion-influenced reactions in which the investigators developed a model for the kinetics of interactions in which memory is significant. The theory is being applied to experiments on excited-state proton dissociation at Tel Aviv University. A second subproject involves development of a theory of kinetic effects when reaction is not necessarily localized on a specific molecule but may be distributed throughout the solution volume. An exact mathematical formulation leads to insuperable analytic problems, but it has proved quite feasible to develop an approximate theory allowing for the estimation of changes in the kinetic behavior.

The classical theory of diffusion influenced reactions uses a model in which only two molecules are available for reaction, whereas in reality the presence of a large number of molecules controls the kinetics. In a joint collaboration with workers in NIDDK members of PSL have developed a theoretical approach to taking many-body effects into account.

Other aspects of the general problem area that have been studied include approximation techniques for solving the Smoluchowski equation and development of a theory of solvent relaxation in ultra-fast laser chemistry.

### Protein Folding

*Byungkook Lee, Ph.D.*
*with H.S. Kang, Ph.D.*

In FY88, PSL embarked on the study of algorithms related to the folding of globular protein molecules. Dr. Kang started by investigating the feasibility of making local dihedral angle motions. Investigators found that perfectly local motions are not feasible but that motion that preserves global orientation and introduces only small global translation is possible. PSL has now written a computer program that incorporates these local motions in the Monte Carlo scheme in the dihedral angle space.

The development of this program is progressing rapidly. It can generate protein structures according to sequential or multi-nucleated mechanisms. The program is based on a simulation technique using predefined probabilities for the dihedral angles found by scanning many known structures. At present, only the hydrophobic force drives the system. Investigators feel that this force is the most important for folding, but a full, successful operation requires including other, more subtle forces including the hydrogen bonding, the van der Waals force, and electrostatic interactions. These forces will be added to research programs in the future.

### Instrumental Analysis

*George H. Weiss, Ph.D.*
*with J.A. Ferretti, Ph.D. (NHLBI); G.S. Spencer, M.D. (Harvard University); U. Shmueli, Ph.D.; S. Rabinovich, Ph.D. (Tel Aviv University, Israel)*

Two projects are included under this heading: the estimation and minimization of error in the design of NMR experiments for chemical analysis, and the calculation of exact representations of probability density functions required for the use of intensity statistics and direct methods for the interpretation of crystallographic data.

In the first of these researchers have analyzed the bias introduced in the estimation of the area under a single peak by the use of apodization filtering. If one does not average

over the range of possible bias, but rather considers the worst possible bias induced by an apodizing filter then it is shown that the error induced by the bias can reach 100 percent or more, depending on the discretizing intervals.

In the area of crystallography PSL has developed the theory required to calculate the probability density functions for crystals with nonsymmetric distributions of structure-factor magnitudes. In addition investigators have begun to examine a variety of possible techniques to correct measurements of low counts that are of the order of magnitude of background noise. The present method simply subtracts the average background that can result in negative counts, which are unphysical. As a further part of the project an introductory monograph on crystallographic statistics will be written.

**Studies in Mathematics and Statistics**

*George H. Weiss, Ph.D.*
*with J. Bendler, Ph.D. (General Electric); D. Ben-Avraham, Ph.D. (Clarkson University); M. Dishon, Ph.D. (NIST); M. Gitterman, Ph.D.; S. Havlin, Ph.D. (Bar-Ilan University); K. Lindenberg, Ph.D. (University of California, San Diego); J. Masoliver, Ph.D. (University of Barcelona); R. Shrager, (LAS/DCRT); K.E. Shuler, Ph.D. (University of California, San Diego); A. Szabo, Ph.D. (NIDDK)*

The two major components of this topic are the development of the theory of diffusive processes in disordered media and the study of model systems for reaction kinetics in chemistry.

A joint project with J. Bendler and M. Dishon involves the generation of accurate tables of the inverse Laplace transform of the function exp(-s-), which is used in a variety of applications in physical chemistry, together with an exploration of approximations that have been suggested in the literature.

A second project is the development of a singular perturbation expansion for the solution of diffusion equations for the kinetics of systems containing reaction centers depending on space but not on time. Analysis assumes that diffusion effects constitute a small perturbation when compared to the distributed reaction rate.

A third project relates to the asymmetry of random walks that are widely used as models for polymer configurations. In an earlier work PSL has shown that the most probable degree of asymmetry corresponds to the maximum possible asymmetry. Researchers have most recently examined how this result is changed by a spatially varying potential field, showing that it can substantially change the symmetry properties of the model.

Together with A. Szabo the staff is writing a monograph on stochastic models for chemical reaction rates.

**Membrane Transport**

*Joshua Zimmerberg, M.D., Ph.D*
*and V.A. Parsegian, Ph.D.*
*with F. Bezanilla, Ph.D.; A. Harris, Ph.D. (John Hopkins University); A. Walter, Ph.D. (Wright State University)*

A large number of specific cellular proteins can be translocated across distinct cellular membranes. Targeting of these proteins is specified by membrane specific signal sequences. It is not known how, after targeting, the polypeptide crosses the membrane. Investigators have studied the permeability properties of membranes fractions enriched in translocation activity by fusing such vesicles to planar phospholipid membranes. We find channel activity, with unitary conductances of 20, 55, 80, and 115 pS in 45 mMK glutamate. A similar large

channel is seen in both dog pancreatic microsomes and E. coli inverted vesicles. Since these two membranes translocate proteins, and glutamate is an amino acid, it is possible that protein translocation occurs through these structures.

The staff has continued to use its newly developed technique of transport-selection to study the gap junction of rat liver. It is generally believed that proteins called connexins form gap junction channels through which neighboring cells exchange cytoplasmic factors. Investigators firmly establish that connexin 32 forms ion channels in single phospholipid membranes, by functional purification and monoclonal antibody identification of the channel-forming protein. Connexin 32 channels show heterogeneous conductances, weak cation selectivity, and asymmetric voltage sensitivity, properties consistent with those expected for a structure that composes one-half the intracellular gap junction channel.

### Cell Membrane Fusion Studies

*Joshua Zimmerberg, M.D., Ph.D.*
*with D. Chandler, Ph.D. (Arizona State University); F.S. Cohen, Ph.D. (Rush Medical College); M. Curran, Ph.D. (NIDDK/LBM); A. Marty, Ph.D. (Ecole Normale Superieure, France); S.J. Morris, Ph.D.; and R. Blumenthal, Ph.D. (NCI/LTB)*

Cortical granule exocytosis in sea urchin eggs is initiated at the point of sperm entry and sweeps over the egg surface 15 to 45 seconds after insemination. This biologically important event releases enzymes and structural proteins that lead to assembly of the fertilization envelope, which is an extracellular matrix preventing polyspermy and protecting the early embryo. The sea urchin egg is an excellent model system for studying the mechanisms of exocytosis. Investigators have taken our previous observations that hyperosmotic sea water inhibits cortical granule exocytosis at different levels depending upon the molecular weight of the osmoticant to the ultrastructural level. The staff has determined by electron microscopy the exact point at which exocytosis is arrested. The researchers find that granule fusion in hyperosmotic sea water is prevented by formation of a granule-free zone at the cortex that separates the cortical granules from the plasma membrane. The few granules that do fuse are arrested at early stages of pore formation that have not been previously visualized. In contrast, polymer solutions do not block fusion. Instead, exocytic pockets are wide open but contain undischarged granule matrix cores. This suggests that polymers retard water entry into the matrix thus preventing proper dispersal of granule components. These observations allow investigators to separate cortical granule exocytosis, normally a single continuous process, into four consecutive but distinct steps.

### Cell-Cell Fusion Due to Influenza Hemagglutinin

*Joshua Zimmerberg, M.D., Ph.D.*
*with R. Blumenthal, Ph.D. (NCI/LTB)*

Initial events in infection by enveloped viruses, such as influenza, rabies, herpes, and HIV, involve binding of viruses to host cell plasma membranes followed by fusion to the plasma membrane or internalized membrane after receptor-mediated endocytosis. PSL is using a recently developed series of realtime fluorescence probe methods to follow the fusion of viruses to cultured cells, or human red blood cells (RBC), to cells expressing viral spike glycoproteins (SGC's) on their extraplasmic

surface. Results of kinetic analysis of these interactions suggest that the membrane rearrangements of proteins and lipids necessary for fusion first allow lipid exchange followed rapidly by establishment of one or more fusion pores, which allow exchange of soluble molecules. The staff directly images the fluorescent dyes by image enhanced video light microscopy; one can then analyze the spatial redistribution of the fluorescent probe between the fusing cells. Investigators developed methods using low light image enhanced video microscopy of live cells to analyze a series of objectives concerned with early events in viral protein-mediated membrane fusion. The hypothesis is that fusion is established by the viral fusion protein forming a pore that gates the passage of molecules between the RBC and target cell according to size and charge. Researchers hope to establish a molecular "time table" that correlates the movement of membrane and cytoplasmic molecules during the fusion process. Using a combination of patch clamp and videomicroscopy, investigators can also correlate membrane and core probe movements with pore formation.

**Histamine Release from Beige Mouse Mast Cells**

*Joshua Zimmerberg, M.D., Ph.D.*
*with M.J. Curran, Ph.D. (NIDDK/LBM);*
*F.S. Cohen, Ph.D. (Rush Medical College); and*
*D.E. Chandler, Ph.D. (Arizona State University)*

Mast cells of the beige mouse contain large intracellular secretory granules (approximately 4 microns in diameter) with membranes that fuse with the plasma membrane in a process called exocytosis. During membrane fusion an exocytotic pore forms that connects the granule interior with the extracellular medium. Through the exocytotic pore the granule contents are released extracellularly and are free to diffuse to target cells.

Researchers have used correlated electrophysiological and light microscopic data to investigate the structure of this exocytotic pore during secretion in mast cells from beige mice. The time course for the widening of the pore is highly variable; it can widen quickly or slowly and can fluctuate between dilated and contracted states of variable conductance (flickering). Initial pore sizes are broadly distributed indicating that such pores are different from traditional membrane channels that have a relatively fixed conductance. Rather, the exocytotic pore is broad with a primary peak between 1 and 4 nano-Siemens, indicating that pore size does not increase in quantal steps. A secondary maximum occurs at about 30 nano-Siemens. Investigators have searched fast-frozen, freeze-fracture replicas of rat mast cells and identified pores with small lumens. The 30 nano-Siemen pores may represent the smallest pores seen in transmission electron microscopy. A model describing fusion on the molecular level that can account for the variable pore sizes, flickering, and known volumes of activation is described.

In a second project, isolated matrices of the giant secretory vesicles of beige mouse mast cells were examined to determine the effects of the ionic composition of the bathing solution on their size. In general multivariant cations condense the matrix relative to univalents.

**Correlation Function Spectroscopy/Neutron and Light Scattering/Biophysical Analysis**

*Ralph J. Nossal, Ph.D.*
*with R. Bonner, Ph.D. (DRS/BEIB); B. Chance, Ph.D. (University of Pennsylvania); S.H. Chen, Ph.D. (MIT); R. Gammon, Ph.D. (University of Maryland); J. Hofrichter, Ph.D. (NIDDK/LCP); V.K. Jain, Ph.D. (DCRT/PSL); S. Krueger, Ph.D. (NIST); J.T. Russell, D.V.M. (NICHD/LNN); G.H. Weiss, Ph.D. (DCRT/PSL)*

A continuing goal is development of methods for studying structural and transport properties of complex biological media. During the past year researchers continued theoretical investigations of various aspects of photon migration in turbid biological media, initiated experimental correlates of those theories, extended the use of dynamic light scattering to probe blood flow in tissue microvasculature, developed methods for using quasielastic light scattering to study the properties of gels and polymer lattices, and performed analyses of neutron scattering measurements of the structure of core constituents of biological vesicles.

A major accomplishment has involved application of random walk theory to time-resolved in vivo optical absorption spectroscopy of biological tissue (with R. Bonner and G. H. Weiss). This work underlies attempts to develop novel noninvasive techniques to measure cerebral blood oxygenation (with B. Chance, et al). Analysis of the time course of the light exiting a tissue after application of a photon pulse has been shown to yield information about tissue scattering and absorption parameters. Preliminary experimental studies (with B. Chance) have verified some of the theoretical inferences, and protocols are being developed to enable investigation of effects of superficial layers (such as the skull) that have different scattering and absorption properties. Investigators also analyzed other aspects of the way that optical path length influences remote optical sensing of biological tissue.

In other work, a mathematical theory of quasi-elastic light scattering (QELS) from particles diffusing within optically dense media has been carried out. Expressions have been derived to interpret features that would be observed if measurements were performed on particles moving within course polymer gels (e.g., a blood clot). QELS has been used to examine the swelling response of neurosecretory vesicles obtained from bovine pituitaries (with V. Jain and J. Russell), the goal is to understand the effects of Ca ions, pH, and other factors that might be involved in the fusion of vesicles to plasma membranes.

In a related study, the staff completed an analysis of neutron scattering measurements made on the core contents of intact neurosecretory vesicles. When released into the blood stream, the core constituents probably circulate as small particles of 22,000 molecular weight, in the form of hormone molecules bound to dimers of neurophysin proteins. Researchers have shown, though, that core constituents exist in much larger aggregates within the vesicles, in the form of particles whose radii significantly exceed 100 angstroms (with S. Krueger, J. Lynn, and J. Russell). Investigators also have analyzed scattering data taken from dense hemoglobin solutions (with S. Krueger, J. Hofrichter, and S. H. Chen), and have established that limited interparticle association probably occurs at high concentrations. This work is being prepared for publication.

**Molecular Forces in Cellular Assembly**

*Adrian Parsegian, Ph.D.*
*with E.A. Evans, Ph.D. (University of British Columbia); D.F. Evans, Ph.D. (University of Minnesota); S.M. Gruner, Ph.D. (Princeton University); R. Podgornik (Lubljana, Yugoslavia, and NIDDK); R.P. Rand (Brock University, Canada); D.C. Rau (NIDDK); J.J. Zimmerberg, M.D., Ph.D. (DCRT/PSL); J.J. Kasianowicz, Ph.D. (NIDDK); K. Gawrisch, Ph.D. (DCRT); and C. Moore (NIDDK).*

By directly measuring the forces between large molecules or between membranes, researchers are learning how these interactions allow the formation of molecular assemblies. Beginning with the discovery that hydration or solvation forces are the most important interactions encountered during the last few nanometers approaching contact, scientists are now systematically investigating the molecular features that govern surface hydration. Hydration forces are very sensitive to relatively small changes in composition of lipid polar groups or of small molecules bound to the macromolecular surface.

Following the measurement of the change in molecular disorder accompanying changes in molecular interaction with separation, investigators derived a model of molecular motion in the force field of neighboring molecules. There has been a connection established between molecular force and motion.

On the theoretical side, researchers analyzed the fluctuations of ionic charge and found that the resulting charge fluctuation forces, analogous to van der Waals forces, create strong attractive forces that explain the puzzling "long range hydrophobic interaction" between nonpolar surfaces. Earlier ideas about this interaction assumed improbable perturbation of water solvent hundreds of angstroms from the surface.

In the study of lipid forces, scientists succeeded in measuring the energies of nonlamellar lipid phases. It appears that many cell membrane lipids that exist in lamellar form are under a great level of bending strain from being forced out of their usual curved conformation. Recognition of this strain has made clear that the incorporation of membrane proteins involves significant energies of protein lipid interaction of bending.

Most recently, investigators have used NMR to detect the influence of membrane lipids on boundary water. It is becoming possible to make a connection between these resonance measurements and the directly measured forces between lipid bilayers.

During this past year the staff has become aware of several features of forces:

- It is possible to describe the packing energy of a large molecule, such as DNA, or a large sheet, such as a lipid bilayer, in terms of configurational entropy of bending while being restricted through direct interactions between neighboring molecules.

- Ionic fluctuations create a van der Waals-like charge fluctuation force. In solutions of sufficiently low ionic strength, these forces take on a range that explains the longest of the puzzling long-range hydrophobic interactions between nonpolar surfaces.

- Phospholipids in the presence of alkanes will often take on nonlamellar inverted hexagonal phases. The preference for this structure indicates a great bending strain in the lamellar structure into which they have been forced.

- It is becoming apparent that NMR measurement of water structuring from deuterium

quadrapolar splitting is practical and that changes in net water structuring compare directly with osmotic stress force measurements.

## Publications and Presentations


Agmon, N.: Viscosity expansions in reactive diffusion processes. *J. Chem. Phys.*, 90:3765–3775, 1989.

Agmon, N.; Weiss, G.H.: Moment approximations for first-passage time problems. *J. Phys. Chem.*, 93:6884–6887, 1989.

Agmon, N.; Weiss, G.H.: Theory of non-Markovian reversible dissociation reactions. *J. Chem. Phys.* (in press).

Ben-Avraham, D.; Weiss, G.H.: Statistical properties of nearest neighbor distances in a diffusion-reaction model. *Phys. Rev. A.* 39:6436–6440, 1989.

Bonner, R.F.; Nossal, R.: Principles of laser flowmetry. In Shepherd, A.P. and Oberg, P. (eds.): *Laser-Doppler Blood Flowmetry.* Kluwer Academic Publ., New York. (in press).

Bonner, R.F.; Nossal, R.; Weiss, G.H.: A random walk theory of time-resolved absorption spectroscopy in tissue. In Chance, B. (ed.): *Photon Migration in Tissues.* Academic Press, New York. (in press).

Chandler, D.E.; Whitaker, M., Zimmerberg, J.: High molecular weight polymers block cortical granule exocytosis in sea urchin eggs at the level of granule matrix disassembly. *J. Cell Biology.* (in press).

den Hollander, W.T.F.; Weiss, G.H.: On the range of a constrained random walk. *J. Appl. Prob.* 25:451–463, 1988.

Ferretti, J.A.; Weiss, G.H.: One-dimensional nuclear Overhauser effects and peak intensity measurements. *Adv. in Enzym.* 176:3–11, 1989.

Gawrisch, K.; Parsegian,V.A.; Rand, R.P.: Membrane hydration. In Glaser, R., and Gingell, D. (eds.): *Biophysics of Cell Surface.* Springer-Verlag Publ., New York. (in press).

Gitterman, S.; Shrager, R.I.; Weiss, G.H.: Influence of noise on systems described by a Mathieu equation. *Phys. Lett. A.* (in press).

Havlin, S.; Kiefer, J.E.; Weiss, G.H.: The trapping problem on a line with dichotomous disorder. *Phys. Rev. B.* 38:4761–4764, 1988.

Havlin, S.; Weiss, G.H.: A new class of long-tailed pausing time densities for the CTRW. *J. Stat. Phys.* (in press).

Horn, R.G.; Israelachvili, J.N.; Marra, J.; Parsegian, V.A.; Rand, R.P.: Comparison of forces measured between phosphatidylcholine bilayers. *Biophysical J.* 54:1184–1185, 1988.

Krueger, S.; Lynn, J.W.; Russel, J.T.; Nossal, R.: Small angle neutron scattering method for in situ studies of the dense cores of biological cells and vesicles: Application to isolated neurosecretory vesicles. *J. Appl. Cryst.* (in press).

Marty, A.; Horn, R.; Tan, Y.P.; Zimmerberg, J.: The delay of the Ca mobilization response to muscarinic stimulation. *Proc. Soc. Gen. Physiol. Series 44.* (in press).

Masoliver, J.; Lindenberg, K.; Weiss, G.H.: A continuous-time generalization of the persistent random walk. *Physica A.* 157:891–898, 1989.

Nossal, R.: An analysis of dynamic light scattering from transient gels. In Abiss, J.B. and Smart, A.E. (eds.): *Photon Correlation Techniques and Applications.* Optical Society of America, Washington, D.C. 132–136, 1989.

Nossal, R.: Quasielastic light scattering from particles diffusing within optically dense media. *J. Phys. Chem.* (in press).

Nossal, R.; Bonner, R.F.; Weiss, G.H.: The influence of path length on remote optical sensing of properties of biological tissue. *Applied Optics.* 28:2238–2244, 1989.

Nossal, R., Kiefer, J., Weiss, G.H., Bonner, R., Taitelbaum, H., Havlin, S.: Photon migration in layered media. *Applied Optics.* 27:3382–3391, 1988.

Oppenheim, I.; Shuler, K.E.; Weiss, G.H.: Stochastic processes in physics. *Encyc. Physics.* (in press).

Pines, E.; Huppert, D.: Geminate recombination and excited state proton transfer reactions: Picosecond dynamics in electrolyte solutions. *Springer-Verlag Publ.*, New York. 48:517–519, 1988.

Podgornik, R.; Rau, D.C.; Parsegian, V.A.: The action of interhelical forces on the organization of DNA double helices: Fluctuation-enhanced decay of electrostatic double-layer and hydration forces. *Macromolecules.* 22:1780–1786, 1989.

Rand, R.P.; Fuller, N.L.; Parsegian, V.A.; Rau, D.C.: Variation in hydration forces between neutral phospholipid bilayers: Evidence for hydration attraction. *Biochemistry.* 27:7711–7722, 1988.

Rand, R.P.; Parsegian, V.A.: Hydration forces between phospholipid bilayers. BBA Biomembranes. (in press).

Shmueli, U.; Rabonovich, S.; Weiss, G.H.: Exact conditional distribution of a three-phase invariant in the space group P1. I.: Derivation and simplification of the Fourier series. *Acta Cryst. A.* 45:361–367, 1989.

Shmueli, U.; Rabinovich, S.; Weiss, G.H.: Exact conditional distribution of a three-phase invariant in the space group P1. II.: Calculations and comparison with the Cochran approximation. *Acta Cryst. A.* 45:367–371, 1989.

Shmueli, U.; Rabinovich, S.; Weiss, G.H.: Exact random-walk models in crystallographic statistics. V.: Nonsymmetrically bounded distributions of structure-factor magnitudes. *Acta Cryst.* (in press).

Shmueli, U.; Weiss, G.H.: Exact random walk models in crystallographic statistics. IV.: PDF's of the absolute value of E allowing for atoms in special positions. *Acta Cryst. A.* 44:413–417, 1988.

Shmueli, U.; Weiss, G.H.: Probabilistic methods in crystal structure analysis. *J. Am. Stat. Soc.* (in press).

Shmueli, U.; Weiss, G.H.; Wilson, A.J.C.: Explicit Fourier representations of nonideal hypercentric PDF's of the absolute value of E. *Acta Cryst. A.* 45:213–217, 1989.

Spencer, G.S.; Ferretti, J.A.; Weiss, G.H.: NMR saturation factors in the presence of chemical exchange. *J. Magn. Res.* 84:223–235, 1989.

Szabo, A., Agmon, N., Zwanzig, R.: Diffusion-controlled reactions with mobile traps. *Phys. Rev. Lett.* 61:2496–2499, 1988.

Taitelbaum, H.; Kopelman, R.; Weiss, G.H.; Havlin, S.: Statistical properties of nearest-neighbor distances at an imperfect trap. *Phys. Rev. A.* (in press).

Taitelbaum, H.; Weiss, G.H.: An approximate theory of photon migration in a two-layer medium. *Appl. Opt.* 28:2245–2249, 1989.

Weiss, G.H.: Biography of Elliot Waters Montroll. *Biographical Memoirs of the Natl. Acad. Sci. USA.* (in press).

Weiss, G.H.: Moviment brownia, difusio, camines aleatoris. *Quaderns.* (in press).

Weiss, G.H.: Simulated annealing. *Encyc. Stat. Sci.* 144–146, 1989.

Weiss, G.H.: Tauberian theorem. *Encyc. Stat. Sci.* 9:184–185, 1988.

Weiss, G.H.: Wald's identity. *Encyc. Stat. Sci.* 9:523–524, 1988.

Weiss, G.H.; Havlin, S.; Matan, O.: Properties of non integer moments in a first passage time problem. *J. Stat. Phys.* 55:435–439, 1989.

Weiss, G.H.; Kiefer, J.E.; Ferretti, J.A.: Accuracy and precision in the estimation of peak areas: The effects of apodization. *Chemometrics and Int. Lab. Syst.* 4:223–229, 1988.

Weiss, G.H.; Kopelman, R.; Havlin, S.: Density of nearest-neighbor distances in diffusion-controlled reactions at a single trap. *Phys. Rev. A.* 39:466–469, 1989.

Weiss, G.H.; Masoliver, J.; Shuler, K.E.: On the asymmetry of a random walk in the presence of a field. *J. Stat. Phys.* (in press).

Weiss, G.H.; Nossal, R.; Bonner, R.F.: Statistics of penetration depth of photons re-emitted from irradiated tissue. *J. Mod. Optics.* 36:349–359, 1989.

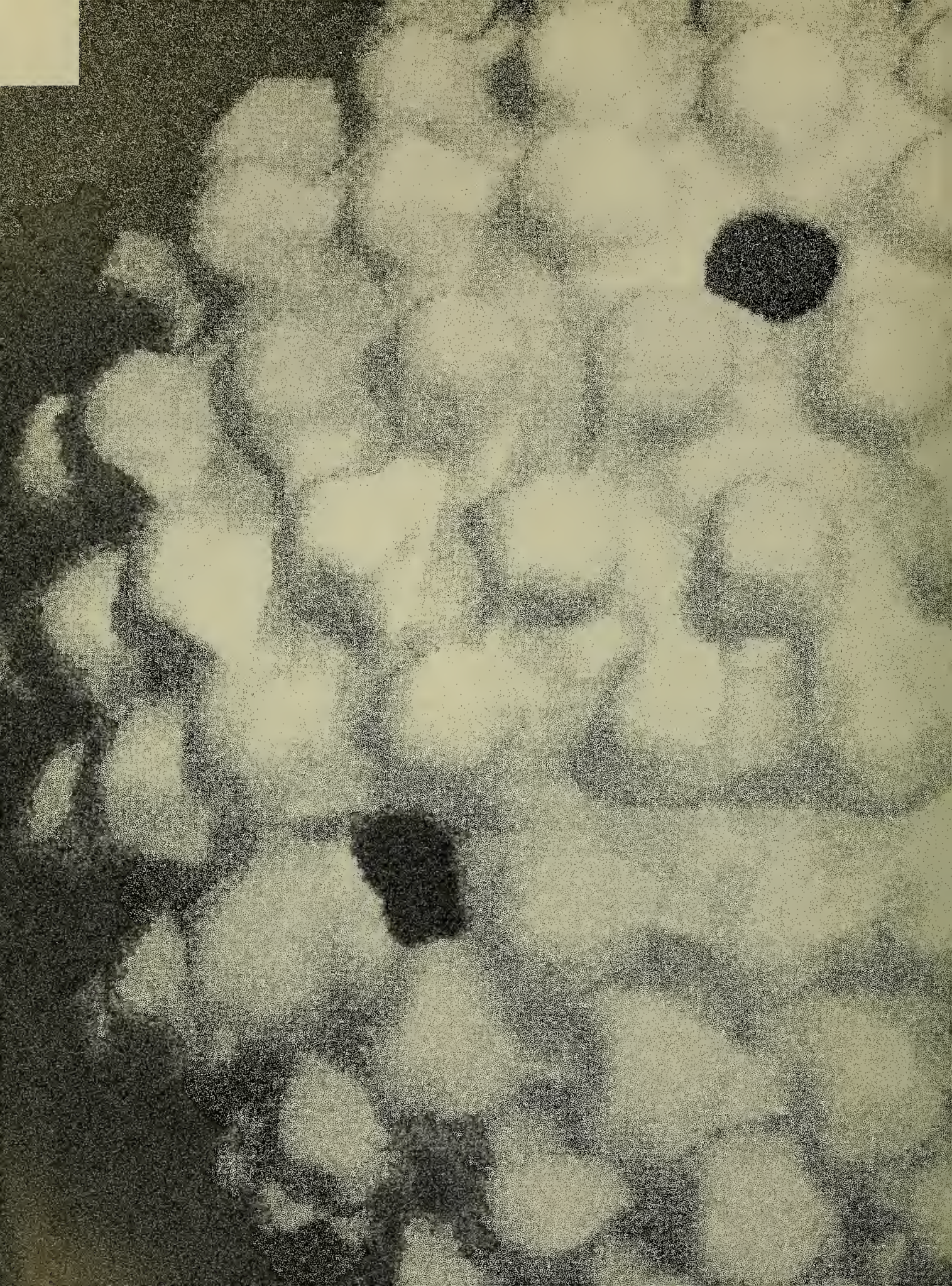
Weiss, G.H.; Nossal, R.; Bonner, R.F.; Kiefer, J.E.; Taitelbaum, H.; Havlin, S.: Random walk theory of photon migration in a turbid medium. In: Jortner, J., Klalter, J. (eds.): *Dynamical Processes in Condensed Molecular Systems.* World Scientific Publ. Co., Singapore, 147–174, 1989.

Weiss, G.H.; Weissman, H.: A measure of the asymmetry of random walks. *J. Stat. Phys.* 52:301–307, 1988.

Weiss, G.H.; Masoliver, J.; Shuler, K.E.: On the asymmetry of a random walk in the presence of a field. *J. Stat. Phys.* (in press).

Weissman, H.; Weiss, G.H.; Havlin, S.: Transport properties of the CTRW with a long-tailed waiting time density. *J. Stat. Phys.* 51:301–317, 1989.

Zimmerberg, J.: Fusion of phospholipid vesicles to planar phospholipid membranes. In: Wilschut, J. and Hoeckstra, D. (eds.): *Cellular Membrane.* (in press).



# Computer Systems Laboratory

*Perry S. Plexico, Acting Chief*

The Computer Systems Laboratory (CSL) develops and applies state-of-the-art computer technologies in areas such as communications, image processing, laboratory automation, and software engineering to NIH's biomedical and clinical research programs. CSL's staff of 31 includes engineers, computer scientists, and computer specialists. The staff also includes people trained in the biomedical disciplines who help to ensure CSL's multidisciplinary capability and act as an interface to the NIH scientific community-at-large.

This year, CSL engaged in 18 projects that span a variety of applications and range in effort from a few labor-months to tens of labor-years. Most projects involve identifying and obtaining commercially available hardware and publicly or commercially available software, then integrating them with specialized hardware or software that CSL designs for particular applications. This usually entails new system- or application-dependent software development, software modification and adaptation, and often, the development of specialized computer interfacing hardware. CSL projects generally require the following characteristics:

- familiarity with the rapidly changing computer hardware and software marketplace to enable selection and purchase of appropriate products and technologies;
- a thorough understanding of computer engineering, science, and technology to resolve issues associated with developing computer hardware and software and with integrating individual components into complex systems;
- a background in biology, medicine, chemistry, or physics to communicate with collaborating NIH scientists.

The laboratory initiates projects when a new computer technology emerges that potentially can benefit the NIH scientific community. Examples include the NIH campus area network backbone and the advanced laboratory workstation projects. Other projects begin with requests for support or collaboration from other BID's to solve particular laboratory computing problems. These projects involve collaboration with a scientist, a group of scientists, or an intramural laboratory who will use the systems or the methods resulting from the project. Examples include the flow cytometry advanced data analysis project, the neuromagnetometer computer system project, and the several projects related to image processing. Finally, CSL provides advice on such topics as computer technology and laboratory computing to the intramural scientific staff, and occasionally, to the extramural staff, academic institutions, and other federal agencies. To make maximum use of available resources, CSL emphasizes projects that potentially have wide impact on biomedical research at the NIH.

## Workstations and Networks Dominate

Workstations and networks continued to dominate CSL activities during FY89, occupying about one-half of CSL's staff. CSL, in concert with the Laboratory Systems Unit, Computer Center Branch, continued the advanced laboratory workstation (ALW) project, to develop a network of powerful, 32-bit, UNIX workstations for scientists to use in their laboratories. The workstations will interconnect via the planned NIH campus area network backbone. A prominent feature of the project is a network distrib-

uted file system that can be shared by workstations throughout the campus and that creates the appearance of a single, large, central file system. In reality, files will be stored on dedicated file servers that also will be connected to the campus network and distributed over the campus.

This year, CSL installed an early version of the Andrew File System, developed by the Carnegie-Mellon University for large scale workstation networks, on a prototype network comprising 5 file servers and 20 workstations. While the system worked well in test use by engineers and computer scientists, the staff did not think it would withstand production use by scientists. The laboratory now has a subsequent release that it believes to be of production quality. CSL also hopes soon to award a contract for up to 15 large capacity file servers. These two steps should enable deployment of the file system for workstation users during FY90.

In addition to the file system work, CSL obtained and installed several major application software packages on the prototype network this year, including MACSYMA, a symbolic algebra system, the SPSS-X and new S statistical packages; Analyze, a medical image rendering package; Prophet II; and the TeX document formatting software. DCRT's Laboratory of Statistical and Mathematical Methodology cooperates in the ALW Project by testing and evaluating some of these software packages.
A closely allied CSL project, the laboratory applications package (LAP) Project, developed a software package to assist in the analysis of laboratory data such as is produced by a variety of spectrometric instruments. LAP already has attracted about 12 scientists in Building 2, who now use it via workstations.

CSL provides network engineering for the NIH Campus Area Network Backbone, cooperates with the Division of Engineering Services in planning and overseeing necessary construction, and will administer and coordinate contractor-provided network installation and implementation as construction progresses. This year, cables were installed between Buildings 12A, 13, and 29; the riser systems in Building 31 were completed; and construction to connect several other buildings was begun. In addition, a contract was awarded for the detailed design, installation, and maintenance of the backbone broadband system.

In May, DCRT recommended the establishment of an NIH Network Policy Board to provide long-term operation, maintenance, and management oversight of the campus network. Concurrent with this recommendation, a DCRT Network Task Group, charged with continued network engineering functions and staffed principally by the CSL network engineering group, was created in DCRT, OD, to give the campus backbone project wider focus and increased stature at this stage of development. The section of this report dealing with the DCRT Office of the Director contains more information about the functions, roles, and responsibilities of the Network Task Group.

## Increased Integration Benefits Projects

This year brought an increased level of integration, cooperation, and technology sharing among CSL's projects. For example, as previously described, the ALW Project will rely on the NIH campus area network backbone. In turn, the ALW Project provides the domain name service,

which translates computer host names to Internet addresses, for all computers on the campus that communicate with the TCP/IP communications protocols. In a similar vein, the flow cytometry advanced data analysis (FC/ADA) project is adapting LAP to its arsenal of computer aided analytical methods.

The FC/ADA staff also perceived the need for a network connection between its development facilities in Building 12A and a flow cytometry laboratory in Building 10. Drawing on networking expertise and cooperation from elsewhere in the division, the project implemented a DECNET connection via a T1 telecommunications link. A member of the FC/ADA staff has agreed to share the experience gained as a result of this work by providing DECNET node name and address coordination on a campus-wide basis under the auspices of the DCRT Connectivity Group.

## Plans for the Coming Year

CSL entered a period of change this year with the retirement of its first and only laboratory chief and one of its three section chiefs. During the long tenure of Alan Demmerle, Chief, CSL and Daniel Syed, Chief, System Design Section, CSL, the laboratory produced many innovative, advanced systems to aid NIH research programs. Technology used for CSL initiatives in the early years of the laboratory, seems primitive compared to projects currently underway. Consider, for example, the contrast between transmitting electrocardiograms over low bandwidth telephone lines in 1970 to today's NIH campus network backbone project, which will accommodate multiple channels, each carrying millions of bits per second. And, the computing power in a single workstation of the type discussed in the ALW project dwarfs the capacity that was shared among many instruments for both data acquisition and processing in CSL's first laboratory automation projects. Demmerle and Syed provided the leadership that, for more than two decades of amazing technological change, let CSL identify new uses for computer technology at NIH and successfully implement a succession of useful projects.

To help address the technological challenges to come, CSL has elected to take this time in its history as an opportunity to conduct a mission review—to examine overtly its past successes and failures and determine its role and functions for the future. An examination of the potential capabilities and applications of modern highly parallel computing systems led the staff to recommend that DCRT implement such a system as an experimental computing resource for NIH science. CSL expects to initiate a procurement action in FY90 for such a highly parallel system. The procurement will include the services of an expert in the techniques of parallel computer programming that will be needed to develop applications for this new class of computer architectures. The acquisition of this new computer will allow the DCRT staff to master these parallel processing techniques that DCRT believes will have significant application in biomedicine. Potential application areas for highly parallel architectures include image processing, biomedical signal processing, and molecular biology sequence analysis.

In FY89, CSL organized a group representing various segments of DCRT to explore ways to improve the division's support for scientific computing. The group recommended that the Division develop a consistent, trans-DCRT UNIX computing environment to form a major com-

ponent of such support. CSL will cooperate with other DCRT labs and branches in the coming year to implement this recommendation.

## Research Projects

### Laboratory Application Package

*J.I. Powell*
*with J. Morris (Brookhaven National Laboratory, formerly with Systex, Inc.)*

The Laboratory Analysis Package (LAP), previously reported as part of the ALW project, is a computer program to assist in the analysis and display of spectrophotometric data. It is an extension of the previous data analysis programs developed by the Computer Systems Laboratory, consolidated and adapted to operate in the UNIX domain. LAP is written in the object-oriented C++ programming language. It was developed on a Sun 3 computer system and has been tested extensively on Sun 3's. Although substantially more powerful than its predecessors, LAP maintains the original philosophy of being an easy-to-learn and easy-to-use, interactive program. LAP is used by the research scientist to manipulate and visualize data collected from laboratory experiments.

LAP can perform a wide range of data manipulations on vector data, x-y paired data, and matrix data using either a command or an expression syntax. Customized command procedures can be saved in files and added to the LAP command set. Results may be viewed as line graphs, scatter graphs, bar graphs, perspective views, or contours in color monochrome viewports within an X window. Results may also be displayed on a Tektronix 4010 compatible terminal.

Following are the major accomplishments of the LAP project this year:

- A preliminary LAP users guide was provided.
- A beta release was installed for the Building 2 pilot study of the advanced laboratory workstation project and is used by approximately 12 scientists.
- The system was ported, but only partially tested, on an IBM RT/PC running AOS.

Next fiscal year CSL expects to complete testing of the IBM RT/PC version, and to develop ports of the software to other UNIX compatible architectures including the DEC VAX, the DECstation 3100, the Sun 4, and the Sun 386i. In addition, the staff expects to make the program easier to use by starting to develop a mouse- and icon-controlled X Window interface for it.

### Hypertension Patient Management System

*S.I. Allen, M.D.*
*with H.R. Keiser, M.D. (NHLBI/IR/OD)*

Investigators are developing a database system to assist NHLBI staff physicians in selecting and modifying drug treatment schedules for hypertensive patients. The methodology focuses on providing staff with rapid and concise information on current and proposed tests, procedures, and cost-effective therapy plans that encourage long-term patient compliance.

Initial personal computer programs built in dBase III+ to store and retrieve available pharmaceutical products, adverse reactions, and drug-drug interactions have been converted to the Oracle relational database. This more powerful tool allows ad hoc queries to be entered in nonprocedural SQL language. Recent efforts have involved the exciting new Hyper-

Card, and subsequently SuperCard, languages as a framework for user-friendly point and click communication with the doctor and data exchange with database modules, and as a means for more rapid program prototyping.

The pilot studies with H.R. Keiser (Chief, Hypertension-Endocrine Branch) involve designing and evaluating explicit queries of the computer database for use by clinic staff. As more experience is gained, automatic queries and checking will be built into a program for use during doctor-patient encounters.

Major steps planned in the future include perfecting hypertension rules and algorithms to optimize patient workups and drug dose titrations in a similar fashion to the prototype Oncocin (tumor) system; simplifying data management functions so that clinincians can use the programs on small economical computers with a minimum of programming assistance; and improving the expert link with the Clinical Center's Medical Information System.

**Auditory Brainstem Response (ABR) Analysis and Interpretation Expert System**

*J.M. DeLeo*
*with A. Pikus, D.A. Sklare (NIDCD)*

The objective of this project is to build an expert system to analyze and interpret the auditory brainstem response (ABR), an electrophysiological response of the brainstem to acoustic stimuli. Expert reading of the ABR is essential in clinical decisions concerning retrocochlear disorders.

Investigators developed a rule-based expert system that performs neurodiagnostic interpretation of the auditory brainstem response. The system was developed on the IBM/PC with the commercial expert system shell GURU and is now operational in the NIH audiology clinic. A talk entitled "An Expert System Approach to Auditory Brainstem Response Interpretation," which describes the expert system, was given at the 1988 American Audiologic Society Meeting in Boston.

ABR interpretation strategies differ considerably among various clinics. This system is a tool for rapidly modeling any ABR interpretative paradigm. Paradigms are easily added, modified, deleted, and selectively engaged at run time. Paradigm rules are Boolean logic expressions containing discrete, continuous, and fuzzy variables representing ABR, subject and stimulus parameters. The system models human ABR expertise in explicit representational form, handles routine cases well, reduces diagnostic errors of omission, and facilitates understanding of ABR clinical decision making. It can be used as a methodology for testing and validating any ABR interpretative strategy, a repository for meaningful ABR interpretative expertise and a medium for sharing ABR knowledge.

Experience in developing and using the system has suggested the possibility of generating more accurate ABR neurodiagnostic interpretation paradigms directly from ABR data using machine-learning methods. Machine-learning refers to the concept of computer programs that systematically generate and evaluate hypotheses for inclusion in the knowledge base. Machine-learning is becoming increasingly important in the attempt to reduce or eliminate the knowledge-engineering bottleneck plaguing most expert system development. Accordingly, the staff has initiated a study to devise programs that derive diagnostic expert system rules directly from ABR data has been initiated. An abstract for a talk entitled "Receiver Operating Characteristic Analysis of Auditory Brainstem

Response Latency" has been submitted for presentation at the 1989 American Auditory Society Meeting in New Orleans. This talk covers CSL's early experience with high-accuracy, data-derived rules for the critical differentiation of retrocochlear and cochlear disorders.

**Neuromagnetometer Computer System**

*R.L. Martino, Ph.D.*
*with C.A. Johnson; T.K. Yap; J.E. Sullivan (DCRT/CSL); S. Sato, M.D.; D.F. Rose, M.D.; P.V. Connaughton; E. Ducla-Soares, Ph.D. (NINCDS/Office of Clinical Director)*

The Medical Neurology Branch, NINCDS, and the Computer Systems Laboratory, DCRT, are collaborating on a research project to noninvasively localize epileptic discharge sources within the human brain by using neuromagnetic recording in conjunction with conventional electroencephalogram (EEG) recording.

Many patients with seizure disorders exhibit low-level cellular discharges between seizures, indicated by interictal spikes or sharp waves in their EEG and magnetoencephalogram (MEG) recordings. This project involves the development of computer techniques for automating and enhancing the procedure that is presently used by NINCDS neurologists to determine the intracranial locations of the sources of epileptiform discharges in patients with epilepsy.

In years previous to this fiscal year, CSL developed an algorithm for the realtime detection of epileptiform discharges from the EEG and MEG signals using the sum of the squares of the first and second derivatives of these signals as a measure of waveform sharpness. This measure is compared to a variable threshold to determine the occurence of spikes or sharp waves. During the past year CSL designed, implemented, and tested a computer system for the automatic detection of these epileptic events in the MEG Laboratory using this algorithm. This system provides a continuous display of the wave forms analyzed indicating where the events have been automatically detected and allows the neurologists to manually select additional events such as longer seizure activity. After the events have been saved by the system, tools have been provided for extracting important parameters from the stored waveform epochs. The extracted waveform values are used to calculate the anatomical location of the sources of the epileptic events.

The development of three-dimensional displays of a patient's head using the digitized outline of the head along with anatomical data obtained from computed tomography and magnetic resonance imaging scans was started during this past year. Neurologists will view these displays to determine where epileptic sources are located relative to actual brain anatomy. An additional 32-bit computer workstation was purchased and is used to perform the electromagnetic field calculations needed to determine the intracranial locations of the sources of the epileptiform discharges.

In the coming year, the realtime detection system will be installed and placed into clinical operation. Work will continue on the development of the three-dimensional anatomical displays and the implementation of the electromagnetic field calculations. CSL will develop methods for classifying the detected epileptiform discharges into types based on their morphology using correlation techniques.

**Advanced Laboratory Workstation Project**

*K. Gorlen*
*with P. Plexico, J. Powell, R. Dew, T. Ghebeles, R. Magnuson, E. Persky, J. Sullivan (DCRT/CSL); J. Dickson, S. Fellini, A. Salemme, J. Small, S. Bailey (DCRT/CCB); S. Orlow (Systex, Inc.)*

The Computer Systems Laboratory and the Laboratory Systems Unit of the Computer Center Branch are working jointly on a project to provide researchers with a network of 32-bit, UNIX-based, high-performance workstations manufactured by a variety of vendors. The most distinctive feature of the network will be the availability of a single file system encompassing all workstations. Potential applications for advanced laboratory workstations (ALW's) include interactive graphics, laboratory data acquisition and analysis, modeling, image processing, and desktop publishing. The staff is adapting software from projects with similar goals that have been underway for several years at Carnegie-Mellon University (C-MU) (Project Andrew) and at the Massachusetts Institute of Technology (Project Athena).

The workstations will be interconnected by the NIH campus-wide network that will provide a means to share resources and access services such as file backup, software maintenance, online documentation, nationwide electronic mail and news, computation and database servers, and laser printers.

Following are the activities of the ALW project this year:

The staff has developed a working prototype system consisting of over 20 Sun-3 and IBM/RT client workstations and 5 file servers with over 8GB of disk storage. Sun-4, Sun-386i, VAXstation 2000, and DECstation 3100 workstations are available and were joined with the network in FY89. The client workstations all run Release 3 of the X Window System and a variety of applications programs. The distributed file system is the Andrew File System (AFS).

CSL issued an RFP for the procurement of up to 15 file servers to be purchased over the next three years and deployed at NIH to provide up to 50GB of storage for ALW clients.

DCRT began participation in the National Andrew File System (NAFS), which is an experiment sponsored by C-MU and IBM to determine the practicality and usefulness of a nationwide distributed file system based on AFS. The ALW staff began production use of a file server backup system developed as part of this project.

CSL installed several major applications, including MACSYMA, a symbolic algebra system, the SPSS-X and New S statistical packages, Analyze, a medical image rendering program; Prophet II; and the TeX document formatting software. LSM has collaborated with CSL by testing and evaluating these applications.

The staff continued its work with the Andrew Tool Kit and began evaluating other X window toolkits such as Open Dialog and InterViews. These toolkits will permit programmers to easily create interactive graphic user interfaces to applications programs.

The staff continued its work with C++ and the development of the NIH Class Library that was used for both LAP (Laboratory Application Package) and the ALW file server backup system. CSL developed the NIH Class Library to supplement C++, a new programming language intended to replace the traditional C language, with a set of reusable Smalltalk-80-like object-oriented programming classes. Members of the staff also wrote a software engineering book entitled, "Data Abstraction and Object-Oriented Programming in C++," that describes this work.

In FY90 CSL plans to begin deployment of AFS servers and advanced laboratory workstations at NIH. To support this, the staff has begun development of a database system to track user and workstation account information. The laboratory also plans to begin integration of the Convex C220, which is replacing the DECsystem-10, with the ALW system.

**Flow Cytometry Advanced Data Analysis**

*L. Barden*
*with R. Tate (DCRT/CSL), S. Sharrow (NCI/DCBD/EIB)*

The flow cytometry advanced data analysis project is a collaboration with the Experimental Immunology Branch, DCBD, NCI, to design and implement a production-oriented research facility capable of archiving and indepth analysis of multiparameter flow cytometry data. CSL will employ a number of complementary analytical techniques, such as multidimensional gated histogramming and nonhierarchical cluster analysis applied to standard format listmode data files (where experimental conditions and parameters are stored in machine readable form along with raw data).

In addition to advanced analytical techniques, such a facility also requires a data staging and archiving system scaled to properly match the expected long-term data acquisition rates and capable of indefinite storage of all raw experimental data. A necessary adjunct to data archiving hardware is an online data base system to provide for user-driven search and access to archived data files.

Network connections are also desirable, both to allow investigators direct access to their data and to provide for alternate means of conducting extensive data analysis and archiving. Early in FY89, a one megabit/second network connection between the development system in Building 12A and the Flow Cytometry Laboratory system in EIB was put in place and has been in steady use since. Software development and data analysis now move easily between the development site in CSL and the EIB laboratory site. An experimental network link to the DCRT Computer Center Branch central facility has been made available to CSL, which will allow direct transfer of large data files so that compute-bound analysis routines can be run on the IBM 3090 vector facility machines.

An initial version of the nonhierarchal cluster analysis program (CAP) is in production use in the flow cytometry laboratory at EIB. The histogram analysis program (LAP) has been ported to run in the FC/ADA environment and is currently in beta test at EIB.

Portions of the data archiving and staging system are being purchased with FY89 funds (a 100 gbyte tape archive and 3 gbytes of online disk storage). CSL is in the process of evaluating magneto-optical disk jukebox systems and file archiving and staging software to complete the archiving system. A second CPU is being purchased for use in the flow cytometry laboratory, along with a directly coupled array processor. This system will be primarily dedicated to executing cluster analysis routines.

EIB is procuring a state-of-the-art flow cytometer, which is expected to become the primary instrument for data acquisition, eventually replacing a Becton-Dickenson FACS-II cytometer (ca. 1976).

**Molecular Graphics, Computer Modeling, and Sequence Analysis**

*B.L. Trus, Ph.D.*
*with A.C. Steven, Ph.D. (NIAMS/LPB); P.M. Steinert, Ph.D. (NCI/DB); B.N. Manjula, Ph.D. (Rockefeller University); S. Havlin, Ph.D.; G.H. Weiss, Ph.D.; B. Fraser (FIC)*

The sequence of some regular proteins, when correlated with other structural information, such as data from x-ray diffraction, fiber diffraction, electron microscopy, and spectroscopic analysis, can be used to evaluate models of protein or polymer structure. Four current studies involve the sequence analysis of keratin and other intermediate filaments (with NIAMS, NCI); sequence analysis of streptococcal proteins (with Rockefeller University); and computer models of biopolymers (with PSL, DCRT), and analysis of protein sequences from viruses and bacteria (with NIAMS, FIC).

As the complete sequence of keratin, other intermediate filaments, and other helical proteins becomes available, an analysis of the sequence can proceed by studying periodicities in the sequence, and by computer prediction of the conformational properties of the specific amino acids in local regions of the chain. These predictions can be used to generalize structures where related sequences are available, and to draw conclusions as to similarities and differences.

Standard Fourier methods have been used to analyze the sequences and to cross-correlate sequences. These sequence regularities are usually correlated with structural features, such as the collagen triple helix, the alpha helix, or the tropomyosin double stranded alpha helix. Using the computer methods of Chow & Fasman or Robson one can predict other types of conformations (e.g. B sheet, coiled, turn).

Additional software has been developed at NIH to illustrate correlations, to create maps of the linear sequences studied, and to view inter- and intra-chain homologies.

This research is significant because many proteins do not form three-dimensional crystalline solids whose structure can be analyzed by classical x-ray diffraction. However, if these proteins are regular, comparison and analogy with related proteins can be used to model the unknown structures in order to understand the structure and functioning of the proteins. In addition, one can use computer models to analyze possible protein structures based on criterion other than regular periodicities.

This year computer models continue to be useful in studying biopolymers. CSL has been studying the amino acid sequence of the Sigma-1 protein from reovirus, as well as images of the protein. Correlation of predicted sequences with observed structure (from image processing of electron micrographs) can yield greater understanding of protein structure and function. Sigma-1, located at the vertices of the icosahedral viron mediates receptor recognition.

As new sequences of regular (helical) proteins become available, it will be relatively easy to model these sequences and describe their structures both graphically and quantitatively.

**Cataract Quantitation Using Image Processing**

*B.L. Trus, Ph.D.*
*with M. Datiles, M.D.; P. Edwards, M.D. (NEI/CB); K. Kashima, M.D. (NEI/CB); M. Unser, Ph.D. (BEIB/DRS); M. Siburg (NEI/CB)*

Images produced by the Scheimpflug principle are being used to quantitate eye opacities in a study to evaluate the potential for the accurate

evaluation of changes in cataract patients. This may provide a means of documenting and monitoring cataracts in vivo, allowing clinical trials of drugs that may prevent or reverse cataract formation.

In cataract research CSL has developed new methods to quantitate, evaluate, classify, and compare slitlamp images of the lens of the eye. New studies are evaluating age-related changes in the normal lens. It turns out that all eyes become more opaque with time, whereas a cataractous lens usually starts with a localized opacity. Researchers are also in the process of evaluating the progression of diagnosed cataracts in follow-up studies, in order to assess the rate at which different types of cataracts develop and the degree of monitoring that is required. Finally, CSL is in the process of comparing images obtained from a Topcon slitlamp camera or Zeiss. In order to detect increases in opacity with time, one needs to use the best diagnostic technique.

Pharmaceuticals are available that may prevent or reverse the cataract formation process. A clinical trial in human patients cannot be pursued because of inadequate means of documenting and monitoring cataracts in vivo. It is hoped that this methodology will provide the statistical and image processing foundation to document and assess changes in lens opacities in cataract patients.

In classification studies the Euclidean distance of a known or unknown from the normal class could indicate quantitatively how typical a normal lens is, or the degree of cataractous change. Another possibility is the differentiation of different groups of cataracts into subgroups according to cause. This method could be used in epidemiological studies, as well as clinical trials, and could provide a basis for diagnosis. In another experiment, investigators are attempting to determine the minimum number of views necessary to characterize different types of cataracts. The Scheimpflug camera can be used to take up to 18 images 10 degrees apart. While two perpendicular views can be used to evaluate some types of symmetrical (isotopic) cataracts, other asymmetric cataracts require six or nine views.

### Viral and Bacterial Structure As Determined by Image Processing of Electron Micrographs

*B.L. Trus, Ph.D.*
*with A.C. Steven, Ph.D. (NIAMS/LPB); M. Unser, Ph.D. (DRS/BEIB); M. Kessel (NIAMS/LPB); B. Fraser (FIC).*

Software is being developed (with NIAMS and BEIB) to analyze viral and bacterial images and perform various statistical and mathematical tests. Such images are typically analyzed by Fourier filtering techniques and the use of correlation alignment together with correlation averaging.

The electron micrographs are taken with a Philips EM400T microscope and the Brookhaven STEM. Some micrographs are preselected by optical diffraction. Negatives are digitized on a Perkin-Elmer 1010MG microdensitometer and analyzed by means of the PIC computer system. Results are photowritten on the Perkin-Elmer microdensitometer and on a Matrix camera station. Image processing uses software developed primarily at NIH.

In virology and bacteriology investigators have published the structure of the cell wall of Bordetella Pertussis, which is composed of a two-dimensional lattice of 40 kDa protein porin. Researchers are currently analyzing three-dimensional tilt sections in order to extend the

understanding of the porin channel's structure to three-dimensions. CSL has been analyzing images of reovirus sigma-1 protein in order to understand the structure/function relation of this molecule. This protein is responsible as the receptor-recognition site for reovirus. In addition, investigators have been analyzing the amino acid sequence of this sigma-1 protein in order to compare the predicted structure with our observation. In the coming year, the staff will begin to analyze electron micrograph images of proteins from the HIV virus with the hope of understanding some structure/function relationships.

Viruses are significantly smaller than bacteria, and as a result are not seen in a light microscope. Information about their structure comes from electron microscopy, which is limited by resolution, low contrast, and noise. If staining is used, the resolution is limited by the size of the stain and often noise as a result of uneven staining. However, because virus structures are generally periodic or contain some symmetry, they are perfect candidates for image processing. This project should be considered as basic research aimed at increasing understanding of the structure and functions of viruses and bacteria in general, as well as of subclasses of viruses similar to those studied to date.

The staff anticipates evaluating other viruses and bacteria for suitability for examination with these methods, and continuing with this ongoing project to determine the structure of various classes of viruses.

**Image Processing of Electron Micrographs**

*B.L. Trus, Ph.D.*
*with A.C. Steven, Ph.D. (NIAMS/LPB); P.M. Steinert, Ph.D. (NCI/DB); M. Unser, Ph.D. (DRS/BEIB); R.J. Podolsky, Ph.D. (NIAMS/LPB)*

This project determines structure from electron microscopy. Suitable software, hardware, and scientific expertise has been provided to allow other scientists, primarily at NIH, to use image processing and computer reconstruction to determine or understand a specimen's structure. Types of data analyzed include intermediate filaments, thin sections of both frozen hydrated myofilament of skeletal muscle and conventional stained embedded muscle, and liver coated vesicles. In addition, researchers are developing computer algorithms to unwarp images, determine resolution, and perform various statistical and mathematical analyses of image data.

The micrographs were taken with a Philips EM400T microscope and the Brookhaven STEM. Some micrographs were preselected by optical diffraction. Negatives were digitized on a Perkin-Elmer 1010MG microdensitometer and analyzed by means of the PIC computer system. Results were photowritten on the Perkin-Elmer microdensitometer or on a Matrix camera station. Images were processed using software developed primarily at NIH.

In the area of image processing algorithms, staff memebers have continued to develop new approaches to iteratively unwarp distorted lattices, and to develop statistical as well as multivariate statistical analysis methods. In addition, a new approach for spectral signal-to-noise resolution development has been devised to process subgroups of average images. This new method is particularly applicable for data sets that contain very large numbers of images.

Finally, new reviews of image processing methods for microscopy have been written and are in press. Investigators will continue to develop new algorithms primarily as problems arise in the anlysis of current imaging problems.

Computer analysis of electron micrographs is still a relatively recent addition to the tools available to scientists for structural analysis. Few laboratories have the combined software and hardware capability to perform the image processing and image reconstruction capability available at NIH. These techniques are especially powerful when applied to two-dimensional crystalline structures. In addition, investigators can correlate and align similar particles that are not crystalline, and correct for a number of artifacts and experimental problems.

Software development will continue as needed. CSL anticipates addition of some three-dimensional image processing capabilities. In addition, as new biological structures become available for analysis, these will be examined.

**Expert Advisor for Shock Treatment**

*D. Syed*
*with K.M. Kempner; H. Fredrickson (DCRT/CSL); J.E. Parrillo, M.D.; W. Hoffman, M.D. (CC/CCM)*

This project has developed, in collaboration with the Critical Care Medicine Department of the Clinical Center, an expert advisor for shock treatment (EAST) for use in an intensive care unit. The drug administration protocol includes the capability for long-term dose maintenance and eventual dose tapering.

Personal computer-based technology (IBM PC/XT or PC/AT) is used in conjunction with an original generated BASIC forward chained inference mechanism.

A knowledge base of approximately 1,400 intervention-specific rules was originally developed for this system. It was supplanted by a revised set that does not explicitly reference interventions (drugs and devices) and is therefore much more adaptable to general treatment concepts. By implementing a table driven mechanism, and using rule iteration, a greatly expanded knowledge base has been encapsulated in approximately 1,000 rules.

After a six-month moratorium due to staff limitations, system development and preparation for formal testing was resumed in November 1988. More rules were added to allow consideration of tentative diagnoses, fluid recommendations were modified to include body surface area, albumin rules were generated, and recommendations for bicarbonate added. Basic rules governing blood gas and pH functions were developed. One of the most extensive and significant modifications incorporated into the system was that of providing and handling alternate nonprotocol recommendations in emergency situations. The explanation capability of the shell was modified to facilitate display of a term-by-term English representation of each rule fired. New explanations were generated, in some cases including line figures. Modifications to the inference mechanism, necessary for the implementation of the system additions, were accomplished. Evaluation methodology and forms were developed, and physician generated scenarios were used extensively to validate the system prior to patient-based trials.

Clinical testing of the prototype system is currently underway and will continue through June 1990. Results of the clinical evaluation will be presented in the future and will allow a final adjustment of the knowledge base.

### Brain Image Registration

*K.M. Kempner*
*with M.V. Green; S.D. Stein (CC/NM); D.E. Rio, Ph.D. (NIAAA/LCS); J.J. Vucich (CC/DR); J.F. Fessler (DRS/BEIB)*

The driving force behind the goal of brain image registration is the need to develop a greater understanding of the processes underlying the generation of PET images. A problem facing researchers involved in the correlation of brain form (structure) from x-ray computed tomography (CT) images and brain function (metabolism) from nuclear medicine positron emission tomography (PET) images. The difficulty concerns the superposition and registration of the tomographic views obtained from these two techniques.

The approach to this problem is based upon a two-stage solution. First, investigators are developing practical methods for the accurate and reproducible placement of the head within a tomographic scanner's aperture. Second, they are developing simplified algorithms for the scaling and registration of digitized images from different scanners on a digital display subsystem.

Precise orientation of the subject's skull within the scanner's aperture is monitored and recorded through the use of a Polhemus Navigation Systems position/orientation transduction subsystem connected to an IBM PC/XT. Image processing and display will be performed using the NIAAA imaging system, consisting of a DEC PDP-11/24 minicomputer and a Gould-DeAnza 6400 Image Processor subsystem.

The development of two inexpensive custom-molded oral appliances allows the position/orientation subsystem's sensor to be fixed to the subject's skull. A targeting algorithm was derived this year to provide the technician operating the system with visual cues related to head position within a scanner's imaging volume. As the subject's head is reoriented, a four-sided polygon on the personal computer's screen translates, rotates, expands, contracts, and its corners re-adjust until a centered square ultimately indicates exact replacement of the skull in a previously scanned position.

Future efforts will center on refinement of the targeting software and in clinical testing of the accuracy and repeatability of skull placement in PET and CT scanners.

### Medical Information Technology Project

*S.I. Allen, M.D.*
*with C.S. Brown, M.D. (Bethesda)*

This project develops better ways for physicians and their associates to use computers in health care recordkeeping for research and patient care. The methodology provides disease-specific and problem-specific protocols and hierarchies of information that allow rapid convergence on relevant diagnoses, treatments, tests, and procedures.

In past years, computer programs were developed for the physician to produce pharmacy prescriptions and drug-related patient information using high-speed menu selection methods. Later, new modules to aid in producing diagnostic schedules and treatment reports were developed. All these programs run on a personal computer (PC), and several PC's may be linked together in a local area network for clinics or practices needing more than one workstation.

Pilot studies with a dermatologist (C.S. Brown) involve the direct entry of medical transactions into a computer terminal by the physicians and staff during the doctor-patient encounter. Initial results with the PC system

show more precise and more rapid prescription writing.

In the current year, pilot programs were rapidly ported to Macintosh hardware with its user-friendly, graphic features and mouse-directed control. The experimental transaction database, maintained in Apple's original HyperCard language, stimulated interest in further exploration of these new tools.

Next year, CSL expects to continue research on the doctor-patient interface using the demonstration dermatology programs on a Macintosh; test object-oriented code and data management techniques written in the new SuperCard (superset of HyperCard) language for ease of update by the doctors; and explore a technology transfer agreement with interested parties in the private sector.

**Zeiss EM902 Electron Microscope**

*W. Gandler*
*with A.C. Steven, M.D. (NIAMS/LPB); B.L. Trus (DCRT/CSL)*

Early this year the Section on Structural Biology, LPB, NIAMS, acquired a Zeiss EM902, the first commercially available electron microscope to incorporate an imaging energy-loss spectrometer. The main thrust of the project—and the reason for procuring this instrument—has been to explore its potential for imaging relatively thick specimens in vitreous ice. Preservation of the native structures of biological macromolecules by rapid freezing and maintaining them in an aqueous environment (ice) during imaging may be one of the most important breakthroughs in electron microscopy of the last twenty years, and the Zeiss EM902 has considerable potential to advance this area of research.

CSL proposed a computer system based on the MicroVAX II to acquire data from the EM902. A Dage-MTI SIT-66 camera provided by Zeiss and incorporated into the microscope is used for image acquisition. Multiple images are averaged for image intensification because the microscope is often operated under minimal dose conditions. The video output from the camera is fed to a frame grabber and image processing board (Data Translation DT 2651 and 2658) connected to the MicroVAX I/O bus. A FORTRAN acquisition program using DT-IRIS subroutines has been developed. One out of every four video frames (CCIR standard) are acquired and averaged in realtime in a 512 X 512 matrix. The averaged image is stored on disk in the same format used by the CSL image processing facility.

Extensive computer image processing can be performed with the highly developed PIC software package on either the large VAX system in DCRT or locally on a MicroVAX III. Either system may be accessed by a network connection. DECNET is used to transfer images on a 56K bit/sec link to the DCRT VAX and a local area VAX cluster connection enables local image transmission.

Shading correction will be applied to the acquired images in the near future.

**DCRT Local Area Network (LAN)**

*R. Fico*
*with L. Freeman, M. Shaffer (DCRT/CSL)*

In order to develop expertise in LAN technology, and provide interconnection of personal computers, terminals, modems and computer systems, CSL developed and installed a small Ethernet network serving DCRT in FY89. Ethernet was selected because of its status as an

established standard and its support by products from a broad spectrum of companies. It was extended to Buildings 12 and 12B in FY85.

Nearly six years after its inception, the DCRT network is heavily used for interoffice communication/terminal-to-host communication, new product evaluation, systems development, and administrative functions. The network has been growing steadily. There are 8 file servers and 15 terminal servers. This year 40 nodes were added, bringing the total to 180 nodes.

In FY85, CSL installed a broadband system to connect the DCRT Ethernet with the Ethernet in Building 13. Baseband to broadband gateways allowed terminal-to-host (DEC-10) connections between these networks. These gateway devices were upgraded to bridges in FY87. As a result of this upgrade, many different types of protocols (e.g., XNS, TCP/IP, DECnet) can now communicate over this link. The significance is that Digital Equipment Corporation (DEC) computers using DECNet, 3Com local office area network equipment using XNS, and powerful workstations using TCP/IP can all communicate over this link.

CSL keeps abreast of network developments and remains active in evaluating promising techniques. In FY87, CSL began evaluating products that interconnect Token Ring systems to Ethernet systems. This involves hardware and software issues that will continue to be investigated.

Evaluation efforts continued in FY88 and were successful with two products that are promising solutions to interfacing token ring networks to the campus wide backbone. The solutions include the 3COM 3+ token ring-to-Ethernet server as one test and a proteon router connecting 3+ networks as another test. The Computer Center Branch of DCRT played a key role in the Proteon Router test.

CSL also evaluated an Ethernet over twisted pair product known as LATTISNET manufactured by Synoptics. Evaluation included a market survey as well as product testing by CSL engineers. This product has gained much attention in the trade press and may serve as an excellent solution to wiring new buildings on the campus.

Evaluation in FY90 will include alternative methods to interconnecting LAN's campus-wide by using the DCRT Ethernet as a base for experimentation. Tests will include routing and bridging products that may serve as network-interface-units between BID LAN's and the NIH campus backbone.

This year, CSL continued working on formal network management procedures. As part of this effort, CSL is developing a data base pertinent to network management. This data base includes information on network configuration, installation requests, quality assurance tests, and remedial maintenance. CSL has also been using a network monitor to collect statistics of network activity and problems. This information is logged on a regular basis. It has been particularly useful in uncovering problems during times when new network equipment is installed. These problems would otherwise have gone undetected and have alerted CSL engineers to persue remedies with vendors when equipment is of questionable performance.

DCRT employees are active in using the DCRT Ethernet LAN and the Computer Systems Laboratory remains an integral part of this by maintaining, understanding, and helping coordinate this activity.

**NIH Campus Wide Network Backbone**

*R. Fico*
*with H. Ostrow, L. Freeman, M. Shaffer, E. Suh, J. Brunetti, and M. Kaeo (DCRT/CSL)*

Starting in FY85, the Computer Systems Laboratory of DCRT began studying the concept of an NIH campus area network. As this project progressed within the CSL, BID networks grew rapidly and the need to interconnect these networks became evident. Today scientific and administrative personnel require interconnection of their independently managed local area networks of computers. DCRT is attempting to meet this need through the implementation of a state-of-the-art communications system known as a backbone.

The backbone consists of a cable system as well as network interface electronics (known as routers) for connecting independently managed local area networks. The cable system contains two technologies: broadband coax and fiber optics. The broadband coaxial technology, identical to that used in the CATV industry, is very mature, supported by the commercial market and used by campuses and large installations nationwide. The broadband technology was emphasized in the early implementation stages of the backbone. However the project is working on a plan to migrate toward fiber and away from broadband as fiber technology matures and standards become available.

A major part of installing the backbone system is the physical construction required to build underground conduit paths between buildings and install cable. The Division of Engineering Services of the Office of Research Services is working on this aspect of the backbone system. Modification within buildings includes the construction of closets to house the cable system electronics and network interface units. Most buildings will have one or two interbuilding closets to connect to the backbone system, however the tall buildings (10, 31, 36, and 37) will have a vertical stack of closets, called risers, for easing access to the backbone system.

In FY89 cables (coax and fiber) were installed between Buildings 12A, 13, 29, and 10. Also in FY89, the Building 31 A, B, and C wing riser systems were completed. The construction to connect Buildings 29A, 30, 36, and 37 is underway. Locations for interbuilding closets were identified in all of the buildings to be connected by the backbone system. With DES help, an architectural and engineering (A&E) firm surveyed all of the proposed building closet locations and determined that all of these locations are physically viable. A&E design work for interbuilding conduit runs and closets for the Group 3 and 4 buildings was completed in FY89.

A contract to design, install, and maintain the backbone broadband system interface electronics was awarded to an experienced company. Under contract, this company began the design of the broadband system in FY89. Some fiber optic work is also being performed under this contract and the fiber links have been terminated and tested between Buildings 12A, 13, 29, and 10.

In FY89, communications between Apollo computers in Buildings 12A and 29 became operational in a test mode using the fiber optic portion of the backbone system. This is a 12-megabit-per-sec. link that has offered network engineers fiber optic design and hands-on experience.

In FY89, DCRT determined that BID LAN's will connect to the campus area network backbone using special purpose network interface

computers known as routers. The network communication protocols to be supported will be TCP/IP and 3COM 3+ (XNS). Attention has also been given to DECnet and details to support this protocol are being addressed.

In order to meet these specialized communications needs for NIH, DCRT formed the Network Task Group (NTG) in May 1989. The NTG is an official detail under the DCRT director, and its formation signifies the attention needed to support networks campus wide. The goals of the NTG are to (1) continue to the backbone system started in DCRT's Computer Systems Laboratory and derive a plan to support this implementation, (2) provide support for BID locally managed LAN's, and (3) provide support of the DCRT local area network.

In FY90, the A&E design work for the Group 5 and 6 buildings will take place and construction contracts for the Group 3, 4, 5, and 6 building groups are expected to occur. Also in FY90, the NTG will determine alternate solutions for long-range support of the campus area network.

### Nuclear Medicine Computer System

*H. Ostrow*
*with S. Bacharach, Ph.D. (CC/NM); M. Green, M. Feldman (CC/NM); D. Vincent (DCRT/DMB)*

CSL has continued to support the Nuclear Medicine Department but on a much reduced level due to demand for resources on other projects. However, CSL continues to consult closely with Nuclear Medicine to help assess their data processing needs and recommend appropriate solutions.

CSL continued to work with DCRT's Data Management Branch (DMB) to identify database requirements for Nuclear Medicine's unique research needs. This year CSL collaborated with DMB to provide a demonstration using Nuclear Medicine's data on a commercial system to illustrate data base management capabilities and provide a focus and direction for the project.

## Publications and Presentations


Douglas, M.A.; Trus, B.L.: An introduction to image processing in medical microscopy, *Medical Progress Through Technology* (in press).

Hagins, W.A.; Ross, P.D.; Tate, R.L.; and Yoshikami, S.: Transduction heats in retinal rods: Tests of the role of cGMP by pyroelectric calorimetry, *Proc. Natl. Acad. Sci.* 1224–1228, 1989.

Kashima, K.; Datiles, M.B.; Trus, B.L.; Edwards, P.A.; Kaiser-Kupfer, M.I.: Localization of lens opacities in gyrate atrophy: Image analysis study with scheimpflug camera, *Folia Ophthalmologica Japonica* (in press).

Steven, A.C.; Mack, J.W.; Trus, B.L.; Bisher, M.E.; Steinert, P.M.: Structure and assembly of intermediate filaments: Multifaceted, myosin-like (but non-motile) cytoskeletal polymers. In: Aebi U., Engel J. (Eds.) *Cytoskeletal and extracellular proteins: Structure, interactions, and assembly.* Heidelberg: Springer, 15–26, 1989.

Steven, A.C.; Trus, B.L.; Brennan, M.J.; Bisher, M.E.; Kessel, M.: Native membrane crystals of Bordetella pertussis' anion-selective porin. In: Dickinson H.G., Goodhew P.J., (Eds.) *EUREM 88.* Philadelphia: IOP Publishing, 457–458, 1988.

Trus, B.L.; Steven, A.C.; McDowall, A.W.; Unser, M.; Dubochet, J.; Podolsky, R.J.: Interactions between actin and myosin filaments in skeletal muscle visualized in frozen-hydrated thin sections, *Biophysical Journal* 55:713–724, 1989.

Trus, B.L.; Unser, M.; Steven, A.C.: Effects on image classification by multivariate statistical methods of different normalization procedures and factorial representations. In: Dickinson H.G., Goodhew P.J. (Eds.) *EUREM 88.* Philadelphia: IOP Publishing, 367–368, 1988.

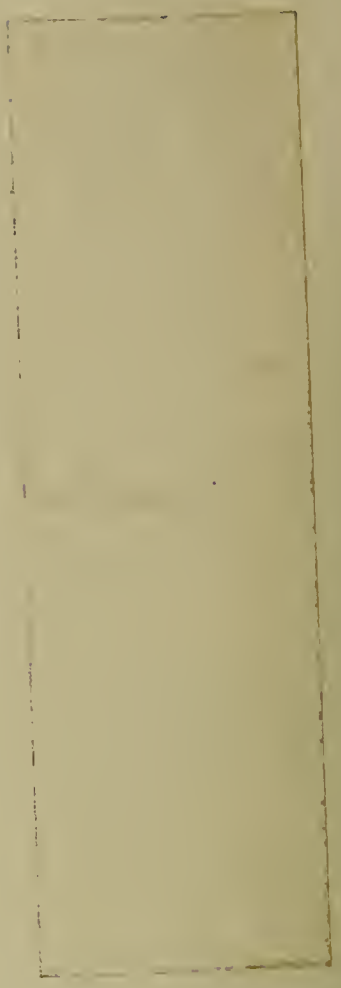
Unser, M.; Trus, B.L.; Eden, M.: Iterative restoration of noisy elastically distorted periodic images, *Signal Proc.* (in press).

Unser, M.; Trus, B.L.; Frank, J.; Steven, A.C.: The spectral signal-to-noise ratio resolution criterion: Computational efficiency and statistical precision, *Ultramicroscopy* (in press).

Unser, M.; Trus, B.L.; Steven, A.C.: Normalization procedures and factorial representations for classification of correlation-aligned images: A comparative study, *Ultramicroscopy* (in press).

Unser, M.; Trus, B.L.; Steven, A.C.; Eden, M.: Iterative restoration of high resolution micrographs of quasi-periodic lattice structures. In: Dickinson H.G., Goodhew P.J., (Eds.) *EUREM 88.* Philadelphia: IOP Publishing, 211–212, 1988.

**Discrimination Prohibited:** Under provisions of applicable public laws enacted by Congress since 1964, no person in the United States shall, on the grounds of race, color, national origin, handicap, or age, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity (or, on the basis of sex, with respect to any education program or activity) receiving Federal financial assistance. In addition, Executive Order 11141 prohibits discrimination on the basis of age by contractors and subcontractors in the performance of Federal contracts, and Executive Order 11246 states that no federally funded contractor may discriminate against any employee or applicant for employment because of race, color, religion, sex, or national origin. Therefore, the Division of Computer Research and Technology must be operated in compliance with these laws and Executive Orders.



DCRT

Division of Computer Research and Technology
National Institutes of Health
Bethesda, Maryland 20892